VOL.
14
1960

THE PHYSICS AND CHEMISTRY OF SOLIDS

AN INTERNATIONAL JOURNAL

VOLUME 14 1960

Proceedings of the Second Conference on Semiconductor Surfaces

Edited by J. N. ZEMEL

Editor-in-Chief HARVEY BROOKS *Cambridge, Mass.*

Associate Editors H. B. G. CASIMIR *Eindhoven* · G. J. DIENES *Brookhaven* · JACQUES FRIEDEL *Paris*

E. M. LIFSHITZ *Moscow*

Editorial Advisory Board

U.S.A.

E. N. ADAMS, Pittsburgh
D. ALPERT, Illinois
L. APKER, Schenectady
J. BARDEEN, Illinois
D. S. BILLINGTON, Oak Ridge
R. G. BRECKENRIDGE, Cleveland
LEO BREWER, Berkeley, Calif.
E. BURSTEIN, Penn.
J. A. BURTON, New Jersey
N. CABRERA, Virginia
G. C. DANIELSON, Iowa
I. ESTERMANN, Washington, D.C.
J. C. FISHER, Schenectady
F. HERMAN, Princeton
CONYERS HERRING, New Jersey
A. R. VON HIPPEL, Cambridge, Mass
J. H. HOLLOMON, Schenectady
W. V. HOUSTON, Houston
H. B. HUNTINGTON, Troy
H. M. JAMES, Purdue
C. KITTEL, Berkeley, Calif.
W. KOHN, Pittsburgh
J. A. KRUMHANSL, Cleveland
A. W. LAWSON, Chicago
H. W. LEVERENZ, Princeton
O. T. MARZKE, Pittsburgh
B. T. MATTHIAS, New Jersey
L. R. MAXWELL, Silver Spring
JOSEPH E. MAYER, Chicago
E. W. MONTROLL, Maryland
DAVID PINES, Princeton
GEORGE T. RADO, Washington, D.C.
H. REISS, New Jersey
J. R. REITZ, Cleveland
A. ROSE, Princeton
O. C. SIMPSON, Argonne
R. SMOLUCHOWSKI, Pittsburgh
F. H. SPEDDING, Iowa
D. TURNBULL, Schenectady
J. H. VAN VLECK, Harvard
B. E. WARREN, Cambridge, Mass.
EUGENE P. WIGNER, Princeton
B. H. ZIMM, Schenectady

Canada

R. E. BURGESS, Vancouver
D. K. C. MACDONALD, Ottawa
W. B. PEARSON, Ottawa

United Kingdom

A. CHARLESBY, Shrivenham, Berks.
W. J. DUNNING, Bristol
F. C. FRANK, Bristol
H. K. HENISCH, Reading
H. JONES, London
J. W. MITCHELL, Bristol
A. B. PIPPARD, Cambridge
M. H. L. PRYCE, Bristol
R. A. SMITH, Malvern
K. W. H. STEVENS, Nottingham
F. C. TOMPKINS, London

France

E. F. BERTAUT, Grenoble
A. GUINIER, Orsay
LOUIS NÉEL, Grenoble
LOUIS WEIL, Grenoble

Germany

M. BORN, Bad Pyrmont
ULRICH DEHLINGER, Stuttgart
R. HILSCH, Göttingen
G. LEIBFRIED, Aachen
E. MOLLWO, Erlangen
H. PICK, Stuttgart
M. VON LAUE, Berlin
H. WELKER, Erlangen

Italy

F. G. FUMI, Pavia

Czechoslovakia

J. TAUC, Praha

Japan

TOSHINOSUKE MUTO, Tokyo

The Netherlands

A. MICHELS, Amsterdam
G. W. RATHENAU, Amsterdam

Sweden

J. O. LINDE, Stockholm
PER-OLOV LÖWDIN, Uppsala

Switzerland

G. BUSCH, Zurich
P. SCHERRER, Zurich

U.S.S.R.

G. S. ZHDANOV, Moscow

PERGAMON PRESS NEW YORK · LONDON · PARIS · LOS ANGELES

STEERING COMMITTEE:

J. N. ZEMEL, Chairman
R. GREENE, Program Chairman
H. SORROWS
P. ROBINSON
N. J. HARRICK
H. HAGSTRUM
F. B. ISAKSON
J. LANDER
M. GREEN
P. HANDLER
D. FRANKL
E. J. SCOTT, Conference Secretary
P. B. WEISZ

ADVISORY COMMITTEE:

W. BRATTAIN
H. E. FARNSWORTH
N. CABRERA
H. STATZ
R. L. PETRITZ

VOL.
14
1960

VO
1
19

CONTENTS

Preface

W. H. BRATTAIN: Introductory Remarks

SECTION 1. CLEANED SURFACES OF GERMANIUM AND SILICON

Chairman: J. A. DILLON, Jr. (for H. E. FARNSWORTH)

Invited Paper

P. HANDLER: Energy level diagrams for germanium and silicon surfaces 1

Contributed Papers

J. T. LAW: Surface properties of vacuum cleaned silicon 9
G. GOBELI and F. G. ALLEN: Surface measurements on freshly cleaved silicon *p–n* junctions . 23
D. R. PALMER, S. R. MORRISON and C. E. DAUENBAUGH: Electrical properties of cleaved germanium surfaces 27
H. D. HAGSTRUM: Studies of Auger electrons ejected from germanium by slowly moving positive ions 33
A. KOBAYASHI, Z. ODA, S. KAWAJI, H. ARATA and K. SUGIYAMA: Impurity conduction of cleaned germanium surfaces at low temperatures 37

SECTION II. NEW TECHNIQUES

Chairman: P. HANDLER

Invited Papers

G. K. WALTERS: Applications of magnetic resonance techniques to the study of semiconductor surfaces 43
R. J. KOKES: Chemical techniques for the investigation of surfaces 51
R. P. EISCHENS: Infrared methods applied to surface phenomena 56
N. J. HARRICK: Semiconductor surface properties deduced from free carrier absorption and reflection of infrared radiation 60
T. H. GEBALLE: Measurements of surface transport phenomena 72
L. H. GERMER and C. D. HARTMAN: Structure of monolayers of adsorbed gases 75

SECTION III. CHEMISTRY

Chairman: P. B. WEISZ

Invited Paper

M. GREEN: Some aspects of the chemistry of germanium and silicon surfaces. 77

Contributed Papers

W. W. HARVEY: Chemical configuration of germanium surfaces in aqueous media 82
W. T. ERIKSEN and R. A. CAINES: The electrokinetic behavior of germanium–liquid interfaces 87
K. H. MAXWELL and M. GREEN: Displacement processes on germanium surfaces . . . 94
R. J. ARCHER: Stain films on silicon 104
M. J. SPARNAAIJ: Gas adsorption and surface charge density of germanium surfaces . . . 111
D. BRENNAN, D. O. HAYWARD and B. M. W. TRAPNELL: Calorimetric determination of the heat of adsorption of oxygen on evaporated films of germanium and silicon 117
S. P. WOLSKY and E. J. ZDANUK: The ion bombardment, oxidation and regeneration of germanium surfaces 124
J. R. LIGENZA and W. G. SPITZER: The mechanisms for silicon oxidation in steam and oxygen. 131

VOL. 14 1960

SECTION IV. COMPOUND SEMICONDUCTOR SURFACES

Chairman: R. L. PETRITZ

Invited Paper

J. J. LANDER: Recent work on surface properties of II–VI semiconductors 137

Contributed Papers

J. N. ZEMEL and J. O. VARELA: Surface properties of PbS semiconductors 142
V. G. BHIDE: Surface sensitivity of rectification 150
J. F. DEWALD: The charge distribution at the zinc oxide–electrolyte interface 155
D. HANEMAN: Structure and absorption characteristics of (111) and (11$\bar{1}$) surfaces of InSb cleaned by ion bombardment and annealing 162
H. C. GATOS and M. C. LAVINE: Etching and inhibition of the III–V intermetallic compounds InSb 169
A. J. ROSENBERG: The oxidation of intermetallic compounds—III. The room-temperature oxidation of $A^{III}B^{V}$ compounds 175

SECTION V. ORDINARY SURFACES OF GERMANIUM AND SILICON

Chairman: H. STATZ

Contributed Papers

W. A. ALBERS, Jr. and J. E. THOMAS, Jr.: Observations of surface transverse magnetoresistance effects on *n*-type germanium 181
A. MANY, N. B. GROVER, Y. GOLDSTEIN and E. HARNIK: Surface mobility measurements in germanium 186
E. HARNIK, Y. GOLDSTEIN, N. B. GROVER and A. MANY: Fast surface states in germanium at low temperatures 193
D. H. LINDLEY and P. C. BANBURY: Surface relaxation effects in germanium at reduced temperatures 200
G. RUPPRECHT: Cross section of mid-gap surface states in silicon by pulsed field-effect experiment 208
S. R. MORRISON: The evaluation of germanium surface treatments 214
H. U. HARTEN: The surface recombination on silicon contacting an electrolyte 220

SECTION VI. THEORY

Chairman: T. B. GRIMLEY

Invited Papers

T. B. GRIMLEY: Surface states associated with adsorbed atoms 227
J. KOUTECKÝ: On the theory of surface states 233

Contributed Paper

J. KOUTECKÝ and M. TOMÁŠEK: Contribution to the theory of the Shockley surface states—I. General formulation and the case of zero deformation of the potential on the Goodwin–Artmann model 241

Invited Paper

M. LAX and P. MENGERT: Influence of trapping, diffusion and recombination on carrier concentration fluctuations 248

Contributed Papers

D. C. GAZIS, R. HERMAN and R. F. WALLIS: Surface elastic waves in semiconductors . . 268
A. AMITH: Galvanomagnetic properties of surface layers in semiconductors 271

Invited Paper

R. F. GREENE: Semiconductor surface transport theory 291

INDEX 299

VOL.
14
1960

AUTHOR INDEX

Albers, W. A. Jr. 181
Allen, F. G. 23
Amith, A. 271
Arata, H. 37
Archer, R. J. 104
Banbury, P. C. 200
Bhide, V. G. 150
Brennan, D. 117
Caines, R. 87
Dauenbaugh, C. E. 27
Dewald, J. F. 155
Eischens, R. P. 56
Eriksen, W. 87
Gatos, H. C. 169
Gazis, D. C. 268
Geballe, T. H. 72
Germer, L. H. 75
Gobeli, G. W. 23
Goldstein, Y. 186, 193
Green, M. 77, 94
Greene, R. F. 291
Grimley, T. B. 227
Grover, N. B. 186, 193
Hagstrum, H. D. 33
Handler, P. 1
Haneman, D. 162
Harnik, E. 186, 193
Harrick, N. J. 60
Harten, H. U. 220
Hartman, C. D. 75
Harvey, W. W. 82
Hayward, D. O. 117
Herman, R. 268
Kobayashi, A. 37
Kokes, R. J. 51
Koutecký, J. 233, 241
Lander, J. J. 137
Lavine, M. C. 169
Law, J. T. 9
Lax, M. 248
Ligenza, J. R. 137
Lindley, D. H. 200
Rosenberg, A. J. 175
Rupprecht, G. 208
Sparnaaij, M. J. 111
Spitzer, W. G. 131
Sugiyama, K. 37
Thomas, J. E. Jr. 181
Tomášek, M. 241
Trapnell, B. M. W. 117
Varela, J. O. 142
Wallis, R. F. 268
Walters, G. K. 43
Wolsky, S. P. 124
Zemel, J. N. 142
Zdanuk, E. J. 124

VOL
14
196

VOL.
14
1960

PREFACE

JAY N. ZEMEL

Chairman, Steering Committee

THE Second Conference on Semiconductor Surfaces was held at the U.S. Naval Ordnance Laboratory in Silver Spring, Maryland, on 2, 3 and 4 December, 1959. The purpose of the Conference, like that of its Philadelphia predecessor, was to bring together scientists of various disciplines whose interest centered on semiconductor surfaces. Six sessions were held. An entire session was devoted to newer techniques that have not as yet achieved widespread use in semiconductor surface studies. The session on clean surfaces of germanium and silicon indicated that this ideal was being approached through thermal or ionic etching and by cleaving. The theory of various surface properties such as transport, noise and the origin of surface states was examined in some detail in the review papers on these subjects. Indicative of the shifting emphasis of the field, a session was spent on the surface properties of II–VI and III–V compounds. Surface chemistry and ordinary surfaces came in for their usual share of attention. Of interest is the trend to carry out surface measurements over fairly wide temperature ranges. More detailed analyses of the various surface parameters should be possible with these added measurements.

There will be three publications issued by the Conference, each one serving a different but useful purpose. The first publication was the Program, consisting of the abstracts and papers. Almost 70 per cent of the papers presented were included in the Program which was distributed 3 weeks before the meeting. The second publication of the Conference is this volume of *The Physics and Chemistry of Solids*. Except for some especially pertinent points, the discussion has been deferred to the third publication, a hard-cover book, which will be published by Pergamon Press in a few weeks. The purpose of this division of the Proceedings is to get the widest and most rapid dissemination for these papers while still having the full Proceedings available for those interested in the details of the discussion. In order to insure the rapid publication of these papers, the Editorial Committee made some minor corrections on the manuscripts.

The total attendance was 175, of whom 164 were from the United States and the remaining 11 from six countries overseas. Surprisingly enough, 11 of the 43 papers in the Proceedings are from overseas, although there is not a one-to-one correspondence between delegates and papers.

A number of organizations and individuals played important parts in the successful conduct of the Conference. The Steering Committee would like to acknowledge the role of the U.S. Naval Ordnance Laboratory, which served as both host and co-sponsor for the Conference. Numerous Laboratory personnel contributed significantly to the Conference. Particular mention should be made of the work of Mr. A. M. ELKSNIS for publishing the 400-page Program in 2 weeks; Dr. F. STERN and the Solid State Division office staff for their co-operation on arrangements; Dr. R. F. GREENE, who was Treasurer and Program Chairman, for his work on arrangements; and Dr. E. J. SCOTT, Conference Secretary, for organizing the papers and discussion.

Substantial financial support, so necessary for topical conferences of this type, was provided by the Naval Ordnance Laboratory and the Office of Naval Research. The contributions of a number of corporations are also gratefully acknowledged.

Finally, the Steering Committee would like to express its thanks to the authors whose promptness and co-operation made this rapid publication possible.

VOL.
14
1960

INTRODUCTORY REMARKS

WALTER H. BRATTAIN

Bell Telephone Laboratories

Dr. HARTMANN, Dr. ZEMEL, MEMBERS OF THE CONFERENCE:

It's really unnecessary that I'm up here. It looks like the Conference has been very well organized. To date, at least, I haven't heard any beefs. We certainly should all express our gratitude to the Naval Ordnance Laboratory and Dr. ZEMEL for organizing this Conference.

You've all heard me comment on what might be done with study of the surface, the importance of it. I won't bore you by saying these things again, but I would like to make a few remarks about the papers to be presented. The papers are interesting. I have at least read most of the abstracts and looked in more detail at some of the papers. It disappoints me at this state of the art to see so many papers in which either one is working solely with a surface in which one can only measure the physics of the surface without any knowledge of what's going on chemically, or at the other extreme one is working with a semiconductor surface with which the physicist could not possibly tell you what the surface was like, or even what the body properties were like, paying attention solely to the chemistry. The hope of this field, of course, is that somewhere along the line we'll learn to do both the chemistry and the physics on the same surface at the same time and from the results reach a new level of understanding of surface phenomena.

There are, of course, bright spots that one can see. Some of us are devising new techniques. People are beginning to make better use of the known techniques as well as devising newer and better ways of getting at what's going on on the surface. Another bright spot is that maybe we are approaching the ideal situation where we have a clean surface to start with. May I suggest that for a semiconductor one needs a new definition of what is meant by a clean surface. In the old days of thermionics a clean surface was something that had not more than at most say a hundredth of a monolayer of foreign material on the surface. This would be quite clean thermionically. I'd just like to make the point that if you take that hundredth of a monolayer and allow it to diffuse into a depth of approximately 10^{-4} cm that you have in most cases a nearly degenerate semiconductor surface, and that this is not what we would like to define as a clean semiconductor surface. What we'd like to have in a clean semiconductor surface is a surface in which the chemical constitution is uniform, or homogeneous, from say the second layer of atoms into the interior, and at the same time not much more than a hundredth of a monolayer on the surface. But that's one of the bright spots of the Conference. We are at least aware of the need for such a definition. How close we are to such a surface I will leave for you to decide after you have heard the various papers on this phase of the work. I think now I'd better close and let all of you get on with the Conference. May I wish all the speakers the best of luck. May they convince you if they can.

VO
1
19

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 1–8. Printed in Great Britain.

SECTION I. CLEANED SURFACES OF GERMANIUM AND SILICON

ENERGY LEVEL DIAGRAMS FOR GERMANIUM AND SILICON SURFACES*

P. HANDLER

University of Illinois, Urbana, Illinois

VOL. 14 1960

Abstract—The concept of localized atomic orbitals has been used to show how the structure and type of binding of surface atoms to the crystal may affect the energy level diagram of the surface region. In particular it has been shown in a qualitative way that the s or p character of the non-bonding orbitals of the surface atoms may determine the energy of the surface state bands and the type of space charge region at cleaned germanium and silicon surfaces.

The transport data of the p-type space charge region of the cleaned germanium have been used to determine the parameters, N_0, the density of states per cm^2 eV, and $(E_F - E_0)$, the degree of degeneracy of a two-dimensional surface state band. Using the value of these parameters obtained from the transport data and the simple model of a two-dimensional surface state band, it has been shown that the linear temperature dependence of the thermionic emission data of Allen on silicon may be predicted to within a factor of 2 or 3.

1. INTRODUCTION

DURING the past few years a great amount of experimental work has been performed on cleaned silicon and germanium surfaces. The results of these experimental investigations have allowed us to develop a crude qualitative picture. This picture describes the cleaned surface as containing a degenerate p-type space charge region and surface atoms which act as acceptors. It is in fair agreement with most of the past data and also with the data presented in the papers of this Conference. Unfortunately, this picture is somewhat clouded by the results of the group of workers at Bell Telephone Laboratories[1] who have shown that boron from the glasses of the vacuum system may contaminate the semiconductor by introducing a p-type layer at the surface, several microns in depth. While this is quite an important effect, I believe it will not change the qualitative aspects of the picture outlined above. This view is borne out for example by the results on cleaved surfaces presented by PALMER *et al.*[2] in this Conference. Their results on the conductivity and field effect of a cleaved germanium surface are in very good qualitative agreement with the results of HANDLER[3] and FORMAN[4], who cleaned their surfaces by means of ion bombardment and subsequent annealing.

In this paper we will assume that this qualitative model of the surface is correct and we shall discuss the physical structure and energy level diagrams which may be associated with this model. In Section 5 of the paper it will be shown how some parameters of the surface, derived from conductivity experiments, may be used to predict the temperature dependence of the thermionic work function of silicon.

2. STRUCTURE OF THE SURFACE

One of the most important prerequisites for a fundamental understanding of the surface is our knowledge of its structure. Although structure may be inferred by many methods, our principal source of information during the last few years has been the work of FARNSWORTH and his colleagues[5]

* This research was supported by the Office of Naval Research.

at Brown University. To a rough first approximation their results indicate that the surface lattice is very close to what one would expect if the volume structure of the crystal were terminated in a plane. In addition, there are other smaller effects such as the half integral order diffraction beams in the (110) azimuths of the (100) face. These beams have been interpreted by FARNSWORTH as indicating that adjacent rows of atoms in the (110) direction have moved closer together to form single or double bonds. Effects of this nature seem probable on the (100) surface since each surface atom has two non-bonding orbitals available and without some sort of bonding the surface would be in a very unstable state. One should expect the line shape and intensity of the half integral order beams to be different from the integral order beams and this is in fact the case.[5] The half integral order beams are lower in intensity and wider in breadth. It would be interesting to see if further analysis of the data could be used to show that the half integral order beams result from the scattering of electrons from two atoms while the integral orders result from scattering of electrons from one atom. In a similar manner it would be interesting to know whether it is possible to distinguish between a monolayer of oxygen atoms in a regular lattice and the normal lattice of the host crystal. Unfortunately, our theoretical understanding of low energy electron diffraction is still poor. It is hoped that progress will be made in this field since it is so very important for an accurate determination of the true energy level diagram of a surface. In the next section it will be shown why greater sensitivity in the electron diffraction techniques is of such crucial importance in understanding semiconductor surfaces. For the purposes of this discussion, it will be assumed that the surface structure of a cleaned surface differs only very slightly from the normal tetrahedral arrangement.

ENERGY LEVEL DIAGRAM OF THE SURFACE REGION

Our efforts to understand the surface of semiconductors will eventually result in an energy level diagram. In this section we would like to discuss a simple model which may be useful in explaining some of the experimental data. This model requires the existence of two adjacent bands, one filled and one empty and differs from the discussion of GRIMLEY[6] who takes only one band into consideration. When only one band is considered no space charge region can arise, even though surface states may be present. For a single filled band with surface states the charge density remains uniform through the surface and volume, since the volume states now have lower amplitudes near the surface. Therefore in order to account for the experimentally observed space charge regions more than one band must be taken into consideration. In what follows I will assume that the germanium and silicon crystals may both be considered a lattice of atoms having atomic properties.

If we consider the isolated column IV atom, it contains four valence electrons which may be distributed among four atomic orbitals, one s and three p. If spin is included, then there are two s-states and six p-states. The s-states are lower in energy than the p-states. When combining chemically with other atoms the most common configuration is that of covalent tetrahedral coordination. The states of the valence electrons can be described by hybridized sp^3 wave functions which are a linear combination of the atomic s- and p-states. Of the original eight states, four are used in bonding to four neighboring atoms and four are raised in energy and called antibonding states. When the atoms are part of a large lattice, the discrete energy levels are broadened. In diamond type crystals, the four bonding states become part of the valence band, the four antibonding states become part of the conduction band and their difference in energy is described as the forbidden zone or energy gap of the semiconductor. As an example we may consider an atom on the (111) surface. It has three bonds to the crystal and one non-bonding orbital. The latter contains two states, one which would have been part of the valence band and one which would have been part of the conduction band. When the non-bonding orbital is empty these two states become degenerate.

Since the surface atom does not have a fourth bond, it will attempt to find a stable configuration by changing the character of its sp^3 bonds. One type of binding might be three pure p-bonds to the crystal and a non-bonding orbital which would be pure s as is shown in Fig. 1(a). The states of the non-bonding orbital by virtue of their s character

VOL. 14 1960

would have energies much lower than the normal valence band states in the volume of the crystal and therefore would give rise to a band of states either below the valence band or overlapping its bottom edge. Since the band lies below the top of the valence band, it would be filled with electrons from higher lying states. However, as each surface atom receives its fourth valence electron, the double degeneracy of the non-bonding orbital is lifted, so that the state for the fifth electron lies at a higher energy. This situation is very similar to the energy level diagram of a group V donor atom in the volume. When a donor atom is ionized, it may capture an electron from the conduction band with either spin and is therefore described as doubly degenerate. However, once the donor is deionized the coulombic potential well is removed, the binding energy of the next electron will be much smaller and the remaining state will have a much higher energy. The degeneracy is lifted and the binding energy of the last electron is usually an order of magnitude less than that of the first electron. One possible energy level diagram which might result from the pure s character of the non-bonding orbital is shown in Fig. 1(b). The band of states S_1 and S_2 are made up from the atomic states of the non-bonding orbital. S_1 represents the lower set of states and S_2 represents the upper set of states which have been split off. Although the bonds to the crystal have also changed their character, it will be assumed that the change in their energies is small compared to the change in the energies of the non-bonding orbital. It is possible, however, that the change may be large and then a second set of surface states would arise. The maximum of their wave functions would lie somewhat below the surface and this second set might be considered as subsurface states. While the effects of these subsurface states may be very important, it will be assumed in the remainder of this paper that their energies lie within the limits of the valence band proper and may be neglected.

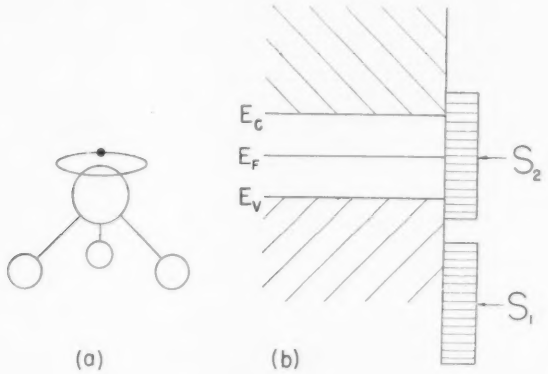


FIG. 1. (a) An atom on a (111) surface containing an s-type non-bonding orbital; (b) an energy level diagram for a (111) surface containing s-type non-bonding orbitals.

A second type of bonding that might be considered is one in which the atom in the (111) surface moves down to become coplanar with its three nearest neighbors as is shown in Fig. 2(a). The bonding to its neighbors would be sp^2 and the non-bonding orbital would be pure p in character and have a much higher energy. Generalizing the problem to that of a crystal again, the pure p-states would be broadened into two bands of states somewhere above the valence band. Fig. 2(b) shows one possible energy level diagram for this situation. Other possible energy level diagrams might show P_1 and P_2 overlapping.

Figures 1 and 2 represent two extreme types of bonding to the crystal. As one goes from a pure s non-bonding orbital through the sp^3 non-bonding orbital to a p non-bonding orbital, the bands of states represented by S_1 and S_2 move up relative

to the valence band until they are represented by P_1 and P_2 of Fig. 2(b).

The energy level diagram of Fig. 1(b) will always give rise to a p-type space charge region while that of Fig. 2(b) may give rise to either an n-type or a p-type space charge region which will depend on the position of the P_1 band with respect to the Fermi energy. The P_1 band is donor in nature while the P_2 band is acceptor in nature. Since the experimental data on germanium and a surface state and a volume state in the free electron model. For example if a state is perturbed out of the valence band, it becomes highly localized and may be considered a true surface state. However, if the perturbation is not large enough to move the state out of its band, the state is still considered part of the valence band but its energy is somewhat shifted. For the purposes of this paper the concept of a surface state will be generalized to include states in the forbidden energy band and

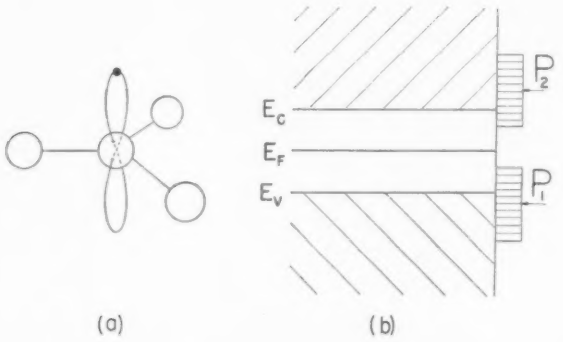


FIG. 2. (a) An atom on a (111) surface containing a p-type non-bonding orbital; (b) an energy level diagram for a (111) surface containing p-type non-bonding orbitals.

silicon indicate the existence of a p-type space charge region and since the coplanar configuration of Fig. 2(a) has not been observed, it seems reasonable to assume that Fig. 1(b) is a better representation than Fig. 2(b). It also indicates that the character of the non-bonding orbital is more s than sp^3. The S_1 band is always filled and its structure may be determined from processes such as the Auger effect. Recently HAGSTRUM[7] has shown that the distribution of electrons ejected from cleaned germanium and silicon surfaces in the Auger process exhibits two peaks. If the energy level diagram of Fig. 1(b) were correct then the higher energy peak would be associated with p or sp^3 localized orbitals which make up the bonds to the crystal and the lower peak would be associated with the s non-bonding orbitals or S_1 band. It should be noted that once the band of surface states overlaps the valence band in energy, there is no longer any clear cut distinction between also those states whose energies have been shifted from their normal values by virtue of their position at the surface. For a discussion of this point, the reader is referred to the paper by GRIMLEY[6] and the references therein.

Although the S_2 band is almost completely empty, its occupation determines the properties of the space charge region. Therefore, by measuring the electronic transport properties of the space charge region, it is possible to determine the properties of the S_2 band in the vicinity of the Fermi energy.

Similar considerations may be used to describe the energy level diagram of a zinc oxide crystal with an excess of zinc present at the surface. If we assume in this case complete ionic binding, the zinc atoms in the volume of the crystal are doubly charged positively. The oxygen atoms are doubly charge negatively, with eight electrons in their outer shell. The eight states (2 s and 6 p) of the

VOL. 14 1960

oxygen atom make up the valence band of the crystal, while the states of the zinc atom are used to make up the states of the conduction band. The lowest lying states of the conduction band are most likely *s* in character. A zinc atom at the surface is not completely surrounded by oxygen atoms and the energy of its *s*-states lies somewhat lower than the *s*-states of zinc atoms in the volume. These surface zinc atoms would then have excess negative charge (be less than doubly charged positively) and it will be assumed that it is two dimensional and circular. Fig. 4 shows an enlarged view of the S_2 band with respect to the valence band. E_0 is the lowest value of the energy in S_2. The wave functions are assumed to be localized normal to the surface, but are free running waves in the directions parallel to it. This statement is only true if the occupation of the band is very small, since S_2 contains only half the normal states of a true energy band in a solid. E_1 represents the highest

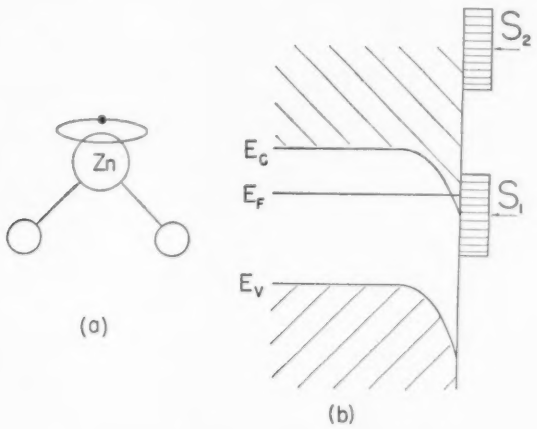


FIG. 3. An energy level diagram for a zinc oxide crystal containing an excess of zinc.

act as donors. If the *s*-states of the surface zinc atoms form a band, their energy level diagram might be like that shown in Fig. 3. The band of states is assumed to overlap the conduction band so that an *n*-type region of space charge is formed at all temperatures. In a similar manner, if there is an excess of oxygen present, a band of states which overlap the valence band would be formed and a depletion region or exhaustion layer would exist at the surface.

In conclusion, we see that in both ionic and covalent crystals, an *n*- or *p*-type space charge region may be formed. The nature of the surface will depend very strongly on its stoichiometry and on the character of its chemical bonds which the surface atom makes to the crystal.

In order to compare the S_2 band of Fig. 1(b) with the results obtained from experimental data, state in the valence band in the space charge region. If the width of the space charge potential well is greater than 100 Å, E_1 may coincide with E_v at the surface. If the width is less than 100 Å, then the uncertainty principle requires that the first levels be somewhat above the bottom of the space charge potential well. This latter condition holds for clean germanium and silicon surfaces and indicates that the density of states as a function of energy in the space charge region may be different from the usual $(\text{energy})^{1/2}$ dependence assumed for the volume.

The narrowness of the space charge potential well is such that the average distance of the holes from the surface region is less than the distance between charges trapped in surface states. For example, in a germanium surface containing 10^{12} electrons/cm^2, the negative charges are about

100 Å apart. The equal number of holes in the space charge region are also approximately 100 Å apart but their mean distance from the surface is less than 25 Å. Therefore a one dimensional diagram cannot adequately describe the region of space charge. In a three-dimensional picture, the holes would move in a potential well made up of a series of overlapping coulombic potentials whose centers are spaced 100 Å apart. No attempt will be made in this paper to discuss this problem further and for the present it will be assumed that the one dimensional diagram of Fig. 4 is adequate.

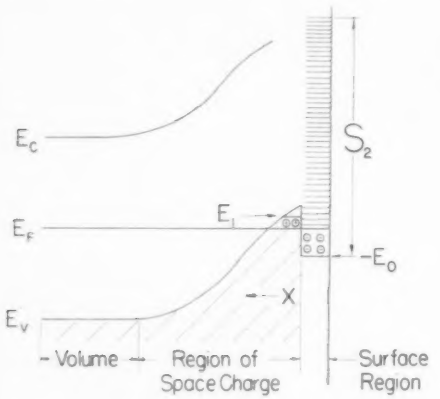


FIG. 4. An energy level diagram for either a cleaned germanium or a cleaned silicon surface. The surface state band is S_2 of Fig. 1.

The density of states per cm² eV for a two-dimensional surface state band is given by

$$N_0 = \frac{4\pi m^*}{h^2} \tag{1}$$

and is independent of the energy. m^* is an effective mass parameter and h is Plank's constant. The number of electrons per cm² in S_2 is just

$$N = N_0 \int_{E_0}^{\infty} f(E)\, dE \tag{2}$$

where $f(E)$ is the Fermi–Dirac distribution function. Integrating equation (2) we have

$$N = N_0 kT \ln \{1+\exp[(E_F-E_0)/kT]\} \tag{3}$$

where E_F is the Fermi energy, k is the Boltzmann constant and T is the absolute temperature. If $(E_F-E_0)/kT > 1$, then equation (3) reduces to

$$N = N_0(E_F-E_0) \tag{4}$$

and N is independent of temperature. Equation (4) is a low temperature approximation. If $(E_F-E_0)/kT < 1$, then

$$N = N_0 kT \ln (2) \tag{5}$$

Thus in the limit of high temperature N is directly proportional to the absolute temperature. If E_0 were greater than E_v, the top edge of the valence band, then

$$(E_F-E_0)/kT < 0$$

and

$$N = N_0 kT \exp[-(E_0-E_F)/kT] \tag{6}$$

In this case, the band of states acts as if it were a discrete state of energy E_0 with a density of N_0kT. It should be noted that most measurements made as a function of temperature are not accurate enough to distinguish the temperature dependence of the prefactor from that of the exponential factor in equation (6).

4. THE REGION OF SPACE CHARGE

One of the most important relationships between the region of space charge and the surface region is derived from the requirement of electrical neutrality. It is that the charge in each region must be equal in magnitude and opposite in sign. If no impurities are present, then for Fig. 4, the number of excess holes in the space charge region, P, must be equal to the number of electrons, N, in the surface state band.

$$N = P \tag{7}$$

Since P and its temperature dependence are known, the quantities N, N_0 and (E_F-E_0) of equation (3) may be determined. P has been found to be equal to a few times 10^{12}/cm² by measurement of the Hall coefficient of a cleaned (100) germanium surface[8] at room temperatures. Also the conductivity as a function of temperature has been shown to be approximately constant over the

VOL. 14 1960

range 77°–300°K.[3] (In the following discussion it must be remembered that most of the data[3, 8] are accurate only to a factor of 2 or 3.) The large value of P requires that the valence band be degenerate or nearly degenerate near the interface between the surface region and the space charge region. It is also expected that the mobility of the holes in this narrow space charge potential well will be relatively independent of temperature. The usual T^{-1} dependence found for carriers in degenerate bands will be absent since the mode of scattering consists of collisions with the surface rather than collisions with the lattice in the form of phonons. If the mobility of the holes is independent of temperature, then P and N are independent of temperature and equation (4) can be used to represent the data. In equation (4) two parameters may be varied. If (E_F-E_0) is made large to satisfy the temperature independence, then N_0 will be small. This implies a very small effective mass for the electrons in the surface state band. Since the mobility of the electrons in the surface state band is usually assumed to be small, N_0 should be chosen as large as possible. Therefore (E_F-E_0) should be chosen small but yet large enough to satisfy the condition $(E_F-E_0)/kT > 1$. It is found that a fair fit of the data can be made with $(E_F-E_0) \simeq 0{\cdot}05$ eV and $N_0 \simeq 4\times 10^{13}$ states/cm^2 eV. Substituting this value of N_0 into equation (1) we find that $m^* \simeq 0{\cdot}1\, m$, where m is the free electron mass. This value of m^* is quite small, but it is the largest value which may be assumed which will be consistent with the temperature independent conductivity and the properties of the model assumed. If this two dimensional model has any validity, it should be useful in predicting other properties of the surface. In the next section it will be shown that the temperature dependence of the thermionic work function of a silicon surface can be predicted from this model using the values of the parameters N_0 and (E_F-E_0) determined from the conductivity experiments. There is of course some question whether the parameters associated with a cleaned germanium surface can be applied to a cleaned silicon surface. Since these parameters are good to only a factor of 2 or 3, this variation may well take into account the actual differences between the germanium and silicon surfaces and the anisotropy of the various crystal faces.

5. TEMPERATURE DEPENDENCE OF THE THERMIONIC WORK FUNCTION OF SILICON

If the two dimensional surface state band exists at a cleaned silicon surface, it will contribute an additional dipole layer to the work function. This dipole layer consists of holes in the valence band and electrons in the surface states. As the temperature is raised electrons will be thermally excited from the valence band to the upper states of the surface state band. These thermally generated electron hole pairs will then add to the dipole layer at the surface. Since the negative end of the dipole is pointing outward, the work function should increase with increasing temperatures. The portion of the work function due to the existence of the surface state band is given by

$$\Delta\phi = qNl/\epsilon_0 \tag{8}$$

where N is defined in equation (3) and l is the mean distance between holes in the space charge region and electrons in the surface state band. Substituting equation (3) in equation (8) gives

$$\Delta\phi = \left\{\frac{qlN_0k}{\epsilon_0}\ln\left(1+\exp[(E_F-E_0)/kT]\right)\right\}T \tag{9}$$

At low temperatures where $(E_F-E_0)/kT > 1$

$$\Delta\phi = \frac{qlN_0}{\epsilon_0}(E_F-E_0) \tag{10}$$

At high temperatures, $(E_F-E_0)/kT < 1$ and equation (9) can be approximated by

$$\Delta\phi = \left[\frac{qlN_0k\ln(2)}{\epsilon_0}\right]T \equiv BT \tag{11}$$

Since all the terms in the square bracket are temperature independent, with the possible exception of l, equation (11) indicates that the work function should be a linear function of the temperature. ALLEN[9] has measured the work function of a cleaned silicon (100) and (110) surface in the range 1300°–1700°K and has found that it may be represented by the following linear function of temperature:

$$\phi = 4{\cdot}30+(3{\cdot}75\times 10^{-4})T \text{ eV} \tag{12}$$

where ϕ is the thermionic work function. Using the values of N_0 and (E_F-E_0) derived from the

conductivity data, it can be shown that B of equation (11) is close to the experimental values of ALLEN if l is assumed to be of the order of 10 Å. At these elevated temperatures l is less than 25 Å, the half width of the well at room temperatures for germanium, and is probably greater than a lattice distance, 5 Å. The value of $(E_F - E_0)$ found from the conductivity data is not large enough for the logarithmic term to be completely independent of temperature. In the temperature range of interest, however, the logarithmic term in equation (9) is equal to about 0·9 and varies about 6 per cent over the temperature range of interest and therefore may be considered constant. Substituting these values into the expression for B in equation (11), we find for $l = 10$ Å

$$B = \frac{qlN_0k(0{\cdot}9)}{\epsilon_0} = 5{\cdot}6 \times 10^{-4}\ \text{eV/°K}$$

Since N_0 and l are known to within only a factor of 2 or 3, this result represents excellent agreement with the data. It would be interesting to see whether contact potential difference measurements from room temperatures up to the range covered by ALLEN's work would be in agreement with equation (9).

6. CONCLUSIONS

The concept of localized atomic orbitals has been used to show how the structure and type of binding of surface atoms to the crystal may affect the energy level diagram of the surface region. In particular it has been shown in a qualitative way that the s or p character of the non-bonding orbitals of the surface atoms may determine the type and energy of the surface state bands at the cleaned germanium and silicon surfaces.

The transport data of the p-type space charge region of germanium have been used to determine the parameters, N_0, the density of states per cm² eV, and $(E_F - E_0)$, the degree of degeneracy of the two dimensional surface state band. Using the value of these parameters and the simple model of a two dimensional surface state band, it has been shown that the linear temperature dependence of the thermionic emission data of ALLEN on silicon may be predicted to within a factor of two or three.

Acknowledgements—I would like to thank Professors MARTIN KARPLUS and JOHN BARDEEN for many interesting discussions.

REFERENCES

1. ALLEN, EISINGER, HAGSTRUM and LAW, *J. Appl. Phys.* **30,** 1563 (1959).
   ALLEN, BUCK and LAW, *J. Appl. Phys.*, to be published.
   J. T. LAW, this Conference, p. 9.
2. PALMER, MORRISON and DAVENBAUGH, this Conference, p. 27.
3. HANDLER P. and PORTNOY W. M., *Phys. Rev.* **116,** 516 (1959).
4. FORMAN R., *Phys. Rev.*, to be published.
5. SCHLIER R. E. and FARNSWORTH H. E., *J. Chem. Phys.* **30,** 917 (1959); *Semiconductor Surface Physics* (edited by KINGSTON R. H.). University of Pennsylvania Press, Philadelphia (1957).
6. GRIMLEY T. B., this Conference, p. 227.
7. HAGSTRUM H. D., *J. Phys. Chem. Solids* **8,** 211 (1959); this Conference, p. 33.
8. MISSMAN R. A. and HANDLER P., *J. Phys. Chem. Solids* **8,** 109 (1959).
9. ALLEN F. G., *J. Phys. Chem. Solids* **8,** 119 (1959).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 9–21. Printed in Great Britain.

SURFACE PROPERTIES OF VACUUM CLEANED SILICON

J. T. LAW

Bell Telephone Laboratories, Murray Hill, New Jersey

VOL. 14 1960

Abstract—Measurement of conductivity, photoconductivity and change in contact potential with light have been made on a sample of floating zone single crystal silicon cleaned by positive ion bombardment and annealed. The absence of a conducting *p*-film after such treatment has been demonstrated. The annealing after bombardment has been measured as a function of time at various temperatures. Changes in the bulk carrier concentration have been studied after various heating treatments.

Changes in conductivity produced by argon ions indicate that in the clean condition the Fermi level is about 0·1–0·2 eV above the valence band edge.

Adsorption of atomic hydrogen decreases the conductivity of both heated and bombarded and annealed samples. At the same time ΔCP becomes more negative indicating a downward movement of the energy bands at the surface.

INTRODUCTION

PREVIOUS attempts to produce clean silicon surfaces have involved the use of either (i) positive ion bombardment[1, 2] or (ii) high temperature heating.[3–5] There is now considerable evidence that under the proper conditions either of these methods will produce a surface which is clean in a two-dimensional sense. By this we mean that there are no adsorbed atoms on the surface, but that there may be a gradation of chemical impurities as one moves into the crystal, normal to the vacuum–solid interface. Such a gradient of chemical impurities near the surface could arise from a depletion mechanism involving the original doping element or from an accumulation mechanism involving a new impurity which reaches the surface from some external source. Some recent experiments[5] have shown that whenever silicon is heated above 1200°K in a Pyrex vacuum system, an increased boron concentration is found in the surface region extending into the bulk some $5–10 \times 10^{-4}$ cm. Obviously then, a surface prepared simply by high temperature heating, while clean in a two-dimensional sense has widely varying properties in the third dimension. We will see later that sufficiently heavy ion bombardment produces a surface which is clean in a three-dimensional sense.

One object of the present work was to obtain and demonstrate the existence of a three dimensionally homogenous surface. In the previous work[5] the existence of a chemical surface film was shown by measuring the conductivity of a thin slice as successively larger amounts were removed from the surface by etching. In the presence of a chemical gradient, the conductivity changes rapidly during the removal of material to a depth corresponding to the penetration of the impurity (usually about 6×10^{-4} cm) and much more slowly thereafter, due to changes in the dimensions of the sample.

One convenient way of both cleaning the surface and removing material in a controlled fashion is by positive ion bombardment. We have used this technique and measured the following quantities to determine the state of the surface during various bombardment and annealing processes: (a) conductivity, (b) photoconductivity and (c) ΔCP, the change in the contact potential of the surface with light. In addition we will describe some changes observed in the bulk properties of the silicon during the heating processes.

The only gas whose adsorption has so far been studied is atomic hydrogen, produced by the thermal decomposition of molecules. Its effect on various electrical properties of both bombarded and heated surfaces will be discussed.

EXPERIMENTAL

The experimental arrangement is shown in Fig. 1. The sample (A) was cut from a p-type Merck, floating zone single crystal with a resistivity of 100 Ω-cm and body lifetime of 200 μsec and etched in CP-8 before mounting in the tube. The ends, A_1 and A_2 were $2{\cdot}2 \times 0{\cdot}7 \times 0{\cdot}9$ cm while the thin center section was $1{\cdot}2 \times 0{\cdot}215 \times 0{\cdot}076$ cm.

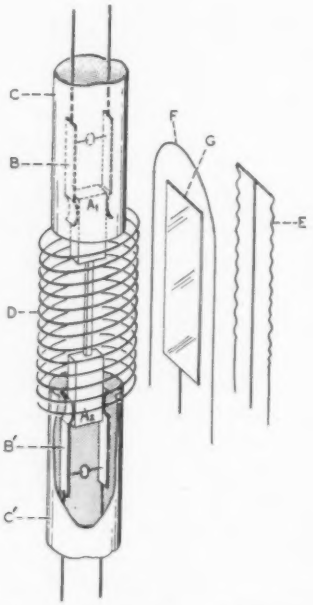


FIG. 1. A schematic arrangement of the experimental tube. The sample (A), surrounded by the cylindrical grid (D), is bombarded by ions from the source (E) (F) (G). The leads and contacts (B) and (B') are shielded from bombardment by the cylinders (C) and (C').

This shape was chosen in order to keep the regions near the contacts cool during the high temperature heating of the thin central section. Previous experiments had shown that even after the molybdenum contacts had been degassed by heating to about 2000°K in vacuum, a subsequent heating to about 1200–1400°K in contact with silicon led to the formation of a chemical p-film immediately around the contact area.

Both contacts, B and B', consisted of two molybdenum springs separated by glass beads so that one pair could be used as current and the other as voltage probes. The silicon immediately under the contact area was rhodium plated to eliminate contact noise. During bombardment we have found it necessary to prevent sputtering of materials other than the silicon. For this reason the contacts were surrounded by cylindrical shields C and C' to which were applied sufficiently large positive voltages that no ion currents to them were observed during the bombardment process.

Bombardment ions were formed by accelerating electrons from the tungsten filament E to the grids F and D in a pressure of argon of 3×10^{-4} mm Hg. The shield G prevented the evaporation of tungsten from the filament to the sample. The grid D was also used as a capacitor plate in the measurement of ΔCP.

The elements shown in Fig. 1 were mounted in a Pyrex envelope with a window for shining chopped light on the sample. The tube was attached to a conventional vacuum system with mercury diffusion pumps and a number of Granville–Phillips valves for admitting gases. After bakeout the pressure was about 2×10^{-10} mm Hg.

The conductivity was measured by a potentiometer circuit and the photoconductivity by chopping the light at 30–110 c/s and displaying the a.c. component of the conductivity on an oscilloscope. The values of ΔCP were obtained by connecting the grid D through an electrometer preamplifier to a wave analyzer. The response of this circuit was calibrated by introducing a standard square-wave voltage at the sample. Unfortunately, for uniform bombardment it is necessary to extend the grid D some distance beyond the ends of the thin central region. Therefore the measured ΔCP signal contains contributions from both the thin central region and the two large end blocks. When all the surfaces being sensed by the grid are homogenous (i.e. after bombardment) this is not too important, but whenever the end blocks are in a different surface condition to the central region, the ΔCP value will have no quantitative significance. In future experiments we intend to change from the present quasi-cylindrical geometry to a parallel plate system where the sensing element will measure only the central region.

RESULTS AND DISCUSSIONS

(a) *The formation and removal of chemical films*

It has recently been shown[5, 6] that whenever silicon is heated above 1300°K in a Pyrex vacuum

VOL. 14 1960

system, a surface layer containing as many as 10^{12}–10^{15} boron atoms/cm^2 is formed. Experiments performed in a quartz vacuum system eliminated this contamination and from this and other experiments we feel that the source of the boron is the Pyrex which contains 12 per cent B_2O_3. Since boron is an acceptor, the surface of the heated sample was strongly p-type with acceptor concentrations at the vacuum–silicon interface as high as 10^{19} cm^{-3} which corresponds to a resistivity of 7×10^{-3} Ω-cm. It was found that the boron doped layer formed even if the Pyrex envelope surrounding the sample were kept at 77°K or 300°K. The vapor pressure of B_2O_3 is given by[7]

$$\log P_{\text{atms}} = 6{\cdot}742 - 16{,}950/T$$

so that even at 700°K it is only 10^{-15} mm Hg. The observed rates of transfer were so high that some other mechanism must have been involved. A series of experiments showed that the rate could be greatly increased by the addition of water vapor to the vacuum system. The reactions between water vapor and B_2O_3 have been studied[8] and the most likely reactions shown to be

$$B_2O_3(s) + 3H_2O(g) \rightarrow 2H_3BO_3(g)$$

$$B_2O_3(s) + H_2O(g) \rightarrow 2HBO_2(g).$$

MARGRAVE[8] measured the rate of removal of B_2O_3 in a stream of helium or nitrogen containing water vapor over the temperature range 553–1173°K and found it to be almost temperature independent. Extrapolating his data to a water pressure of 10^{-2} mm Hg one can obtain an evaporation rate of 3×10^{13} boron atoms sec^{-1} which is easily sufficient to explain our observed changes.

Our model of the boron transfer process is then as follows. Either during the glass blowing operations (when there is a high concentration of water vapor in the system from the torch) or when the system is at or near atmospheric pressure, water vapor reacts with the B_2O_3 in the Pyrex to form the volatile compounds H_3BO_3 and HBO_2. These then diffuse over to the sample surface where subsequent heating decomposes them to B_2O_3 which in turn is reduced by the silicon to elemental boron. This then diffuses in a distance which is determined by the time and temperature of heating.

If this mechanism is correct, the contamination can be avoided by eliminating either the boron or the water vapor from the system. Use of a demountable quartz system proved satisfactory but when a bakeable quartz system was constructed it proved impossible to prevent it being contaminated by boron from the Pyrex diffusion pump. A more feasible way involves the removal of the water vapor. After baking and pumping, the partial pressure of water in a normal high vacuum system is less than 10^{-10} mm Hg so that if the B_2O_3 already deposited on the silicon could be removed, the rate of redeposition should be very low. We have used ion bombardment to accomplish the removal.

After baking the system to obtain pressures near 2×10^{-10} mm Hg and degassing the silicon at 1100°K for several hours, the sample was heated to 1580°K for several minutes by passing current through it. Following this, its room temperature (300°K) conductivity was measured and then a series of ion bombardments carried out with 1000 V argon ions. After each bombardment the silicon was allowed to cool to 300°K and the conductivity remeasured. In Fig. 2, values of $\Delta\sigma$, the change in room temperature conductivity produced by the bombardment in mho/square are shown as a function of the number of incident ions per cm^2 (N_I) in curve (1). For quite small numbers of ions, the conductivity decreases by about 20 μmho/square and then remains constant until an appreciable depth of silicon is removed, after which it once more decreases. We suggest that the first effect is due to the presence of argon ions on the surface together with possibly some shallow bombardment damage and is simply a change in the space charge region. The second larger effect is caused by the removal of a chemical layer of high conductivity. If we assume that the efficiency of the bombardment process is unity,* i.e. each incident ion removes one silicon atom, we can replot the data given in Fig. 2 as $\Delta\sigma$ against the depth of material removed. The result of this is shown in Fig. 3 together with corresponding points (crosses) obtained by controlled etching of a silicon slice which had been similarly heated in vacuum. The excellent agreement between the two sets of data is probably fortuitous as the amount of B_2O_3 transferred to the silicon depends on a number of uncontrolled variables. Nevertheless the agreement shows that

* The sputtering yield of germanium in argon has been determined recently[9] and found to be unity for 450–500 V ions.

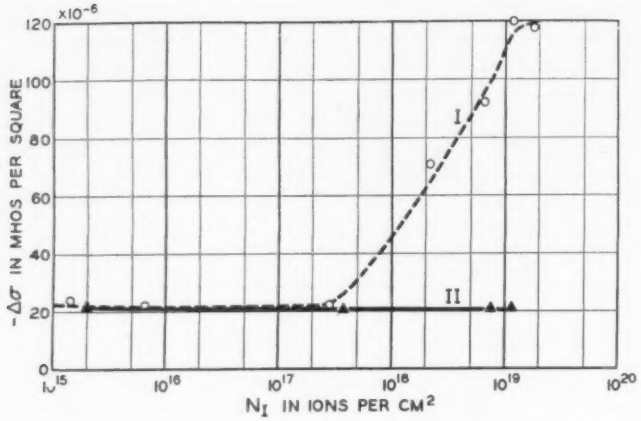


FIG. 2. The decrease in conductivity of the silicon (in mho/square) during ion bombardment as a function of the number of incident ions (N_I). Curve (I) refers to a sample with a heavily boron doped surface film, while curve (II) was obtained from a three dimensionally homogeneous sample.

the major effect of prolonged bombardment is the removal of material from the surface without changing the properties of the underlying silicon.

Following the heavy bombardment shown in Fig. 2, the silicon, was allowed to stand in a vacuum of 2×10^{-10} mm Hg for 3 days. It was then reheated to 1580°K for a total time of 60 min, the heating being carried out in 10 sec flashes with 10 sec intervals between them. During this time the measured bridge conductivity did not change,

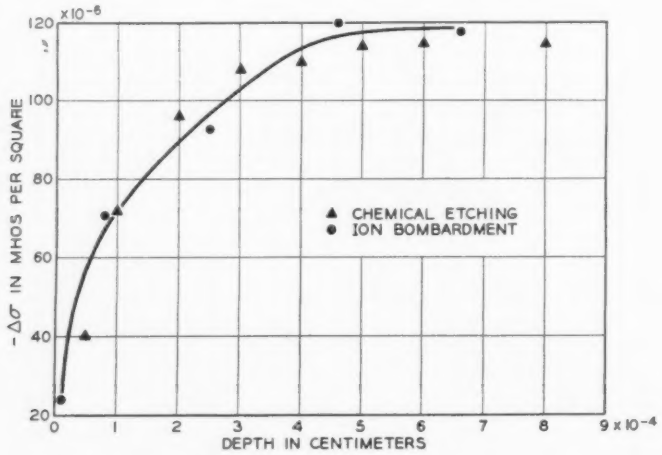


FIG. 3. The decrease in conductivity of the silicon as a function of depth. The circles represent the data shown in Fig. 2, curve (1), assuming unit efficiency for the bombardment process. The triangles are data obtained by chemical etching of a vacuum heated silicon slice.

VOL. 14 1960

but this, however, is not conclusive unless one can show that the bulk conductivity is remaining constant. For this reason the silicon was then rebombarded with argon ions and resulting $\Delta\sigma$–N_I curve obtained is shown in Fig. 2 as curve (2). In this case the initial change in σ at low ion densities was present but the large decrease in conductivity in the micron removal range was absent. This is exactly what one would expect if water vapor is necessary for the boron transfer process. Its partial pressure in a good vacuum is very small and therefore no transfer is possible. These experiments were repeated twice more by letting the vacuum line down to atmospheric pressure for one day, pumping, baking and heating the sample to 1580°K. In both cases the initial bombardments gave curves similar to curve (1), Fig. 2, while subsequent bombardments were identical to curve (2), Fig. 2.

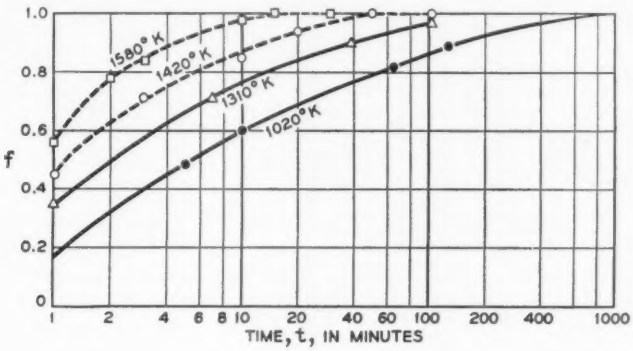


FIG. 4. The fractional change in conductivity of a bombarded surface as a function of annealing time at 1020, 1310, 1420 and 1580°K.

In summary then, we have demonstrated that positive ion bombardment will remove the boron containing surface film on vacuum heated silicon and that it will not be reformed by heating even at 1580°K under good vacuum conditions. Its presence or absence is most easily demonstrated by obtaining curves such as are shown in Fig. 2.

(b) *Annealing positive ion bombardment damage*

If one is to use ion bombardment to remove contaminated layers, then information on the annealing rate of the damaged material left on the surface and the rate of removal of embedded argon atoms is of some importance. The silicon was bombarded ($\sim 10^{19}$ ions cm^{-2}) and then heated to 1580°K for several hours to anneal any damaged material completely. It was then given a series of bombardments (6×10^{18} ions cm^{-2}) and after each one, annealed at a fixed temperature until the conductivity reached a constant value. In this way, curves of conductivity versus annealing time were obtained at 1020, 1310, 1420, 1580°K. The final conductivities for the latter three temperatures were identical but because of a change in bulk conductivity (which will be discussed below) the final value at 1020°K was much higher. This run was then repeated by bombarding after a lengthy heating at 1020°K. The final conductivity value after annealing was then the same as that obtained before the bombardment but almost three times higher than the value found after heating above 1300°K. In Fig. 4, the fraction, f, of donors annealed at each of the four temperatures, is shown as a function of time. At least two effects probably contribute to the annealing process, (i) the removal of embedded argon atoms and ions and (ii) the annealing of surface mechanical damage caused by the ions. By flashing the sample to 1580°K immediately following a bombardment and measuring the pressure change in the system, the number of gaseous species desorbed was found to vary between 1×10^{15} and $3{\cdot}5 \times 10^{15}$ atoms cm^{-2} for values of N_I between 10^{15} and 4×10^{18} ions cm^{-2}.

The results shown in Fig. 4 indicate that we are not dealing with a simple process and no interpretation will be attempted. One difficulty is that the conductivity is affected by any argon ions remaining on the surface as well as the bombardment damage itself. The data, however, show that the time required for annealing to a constant conductivity varies from 15 min at 1580°K to about 700 min at 1020°K. This in fair agreement with DILLON and FARNSWORTH[1], who found it necessary to anneal a sample used for work function measurements at 1263°K, for at least 30 min.

(c) *Thermally caused changes in bulk conductivity and trap density*

Before the sample had been heated above 800°K, it had a bulk resistivity of 100 Ω-cm and a filament lifetime of 50 μsec. Neither of the quantities changed appreciably on heating to 1020 and 1190°K for several minutes, but 1 min at 1310°K caused a four-fold increase in the lifetime together with a small increase in the resistivity. Successive 1-min heatings at 1420, 1480 and 1580°K gave further increases in both the photo decay time and the resistivity. The true bulk value of the resistivity at this time was difficult to determine as the surface had not been bombarded and later experiments showed that it was covered by a highly conducting layer. Hence, the values of resistivity obtained after heating are minimum ones. In Table 1, the observed changes in photo conductivity, half-life and resistivity after various heat treatments on the unbombarded sample are summarized. A heavy bombardment to remove the conducting surface film, followed by an anneal at 1580°K gave a bulk resistivity of 300 Ω-cm but did not change the half-life of injected carriers in the sample. The effect of heat treatment on these properties was then systematically studied after the surface film had been eliminated. The silicon was heated to a given temperature, radiation quenched to 300°K and measured. This process was repeated until no further change occurred after which a different temperature was investigated. In Fig. 5 the resistivity (in Ω-cm) is plotted against the temperature at which the sample was held until the steady state was reached. Above 1300°K the resistivity is constant and equal to about 300 Ω-cm, while below this it steadily decreases. The increase in resistivity corresponds to a decrease, in acceptor concentration (or increase in donor concentration) of about 1×10^{14} cm^{-3}.

Data were not obtained below 1020°K because even at this temperature the heat treatment process was very slow. In Fig. 6 some typical rate curves are shown for 1020, 1100 and 1420°K treatments. The 1020°K curve was obtained on a sample previously stabilized at 1580°K, while the other two curves were each taken after a 4-hr heat at 1020°K. The steady state conductivity obtained for any temperature was independent of whether the previous heat treatment was at a higher or lower temperature.

The apparent lifetime of the sample could similarly be altered by heat treatment from values as high as 3000 μsec at 1580°K to less than 500 μsec

Table 1

Treatment	Photoconductivity (mV)	Half-life (μsec)	Resistivity (Ω-cm)
Initial	0·6	50	100
Bakeout (725°K)	1·0	60	100
1020°K—2 min	0·4	20	100
1190°K—1 min	0·3	15	100
1310°K—1 min	1·0	60	110
1420°K—2 min	8·5	400	130
1480°K—1 min	10·0	—	140
1580°K—1 min	12·0	—	140
1580°K—5 min	16·0	—	140
1580°K—60 min	20·0	1000	140

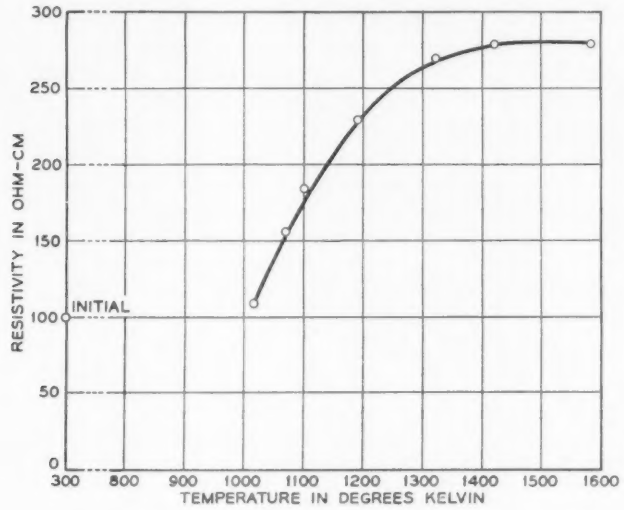


FIG. 5. The bulk resistivity as a function of the heat treatment temperature. The initial value shown (100 Ω-cm) refers to the crystal as grown.

at 1020°K. Considering the thinness of the sample used (0·076 cm) no reasonable values of body lifetime and surface recombination velocity could give such high filament lifetimes. If a d.c. light was shone on the sample during the measurement, the apparent lifetime decreased steadily as the intensity of the additional light was increased. This result is consistent with the picture of a number of temporary trapping centers which can be kept partially filled by the hole–electron pairs created by the d.c. light. The existence of such centers in silicon has been previously investigated

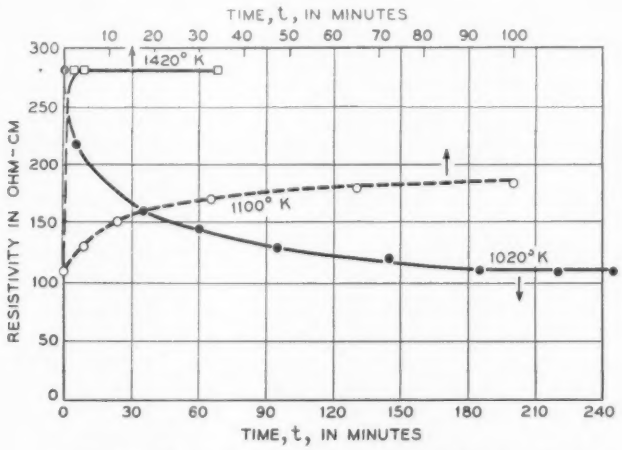


FIG. 6. The rate of change of the bulk resistivity during heat treatment at 1020, 1100 and 1420°K. The curve for 1020°K refers to the lower time scale, the other two to the upper.

by HORNBECK and HAYNES[10, 11] and the traps which we have observed appear similar to their shallow traps which had lifetimes in the millisecond range. We have looked for, but failed to find, any deep traps which would have appreciably longer time constants.

There are at least three possible chemical origins of the observed changes in carrier and trap density, (i) oxygen, (ii) vacancies and (iii) some donor of low solubility.

(i) It is known[12] that oxygen in silicon can act as a donor after suitable heat treatment and that the donors can be annihilated by changing the heat treatment conditions. It has also been suggested[13] that the trapping centers in silicon are associated with oxygen in the crystal.

It seemed unlikely that oxygen was the cause of the changes found in the present work because the starting material, floating zone silicon, is known to have a very low oxygen concentration.[14] In addition, the donor concentration is proportional to the fourth power of the oxygen concentration[15] so that we would require about 10^{17}–10^{18} oxygen atoms cm^{-3} in the starting material. These could not have diffused in from the silicon surface as the time required to produce the increase in resistivity at 1300–1400°K was much too short. Nevertheless a 16-hr. heat treatment was carried out at 720°K (a temperature at which oxygen donor formation is known to occur[12, 16] but no changes in either donor concentration or trap density were observed.

(ii) A second possible mechanism is the diffusion of vacancies into the lattice from the surface where they might be formed during the evaporation of silicon at high temperatures. This would not explain the changes observed initially but could be a contributing factor during subsequent high temperature heating. However, as Fig. 5 shows, the resistivity is unchanged in going from 1420 to 1580°K while over this same temperature range the evaporation rate[17] increases from $1{\cdot}5 \times 10^{15}$ to 4×10^{16} atoms cm^{-2} sec^{-1}. It is therefore unlikely that the observed changes are due to the introduction of vacancies.

(iii) The third, and what appears to be the most likely explanation, is that we are dealing with the precipitation and redissolving of a donor impurity. If the crystal, when grown, contained a donor of low solubility, it would probably be present in a precipitated form due to the slow cooling rates used in crystal growth. On subsequent heating to about 1300°K some of the precipitated impurity could dissolve and remain in solution after the radiation quench to 300°K when it would appear as an electrically active donor. We can neglect the diffusion of such an impurity in from the surface for the same reasons as were described above in section (i), for the case of oxygen. Such precipitation processes have been investigated for only a few elements in silicon[18, 19] and germanium[18, 20] and from the present data it is not possible to decide which element or elements are involved.

It is known[21] that the transition elements, iron, cobalt and manganese, have low solubilities ranging from about 10^{16} atoms cm^{-3} at 1450°K to about 2×10^{15} atoms cm^{-3} at 1200°K for iron, and 1350°K for cobalt and manganese. Both iron and manganese also act as donors with energy levels near the middle of the band gap. In the present sample the Fermi level is also near the middle of the gap so that only a small fraction of the donor states will be ionized at room temperature. Hence, the number of conduction electrons measured will be much less than the total number of dissolved impurity atoms.

In the case of germanium, it has been found that deep lying levels due to iron group impurities can act as trapping centers below room temperature.[22] Similar information on silicon is lacking, presumably because the low solubility of such elements makes it difficult to determine details of their action as trapping centers.

The principal disadvantage of these deep lying levels in attempts to study the surface properties of silicon lies in the high trap density which they produce. Measurements of surface recombination velocity are then almost impossible as the photo signal caused by the recombination of carriers is only a very small fraction of that due to the traps. From our observations, it would appear that the difficulty can be avoided by never heating the sample above 1200°K. We have found that this is sufficiently high to anneal out the bombardment damage and FARNSWORTH *et al.*[1, 2] have shown that it is suitable for initial crystal degassing.

(d) *Effect of positive ion bombardment on surface properties*

A number of studies have been made on the effect of ion bombardment on both silicon and

Table 2. Effect of ion bombardment on Ge *and* Si *surfaces*

	Ge	Si
Depth of damage for 80/kV ions	$1{\cdot}3\times10^{-5}$ cm[23]	$1{\cdot}0\times10^{-4}$cm[24]
Majority carrier type at surface	p on either n- or p-type[23] substrate	n on p-type substrate[24] p on n-type substrate
Surface recombination velocity	$>10^4$[25]	$\sim5\times10^3$[24]
Minimum ion energy to produce rectifying surface layer	20 keV[23]	~1 keV[26]
Minimum annealing temperature	773°K[23]	1073°K[24]

germanium surfaces. The results are summarized in Table 2.

When we had made sure that there were no chemical inhomogeneities in the surface region due to boron doped layers, we carried out a study of the effect of ion bombardment on the conductivity, photo conductivity and the change in contact potential with light. The latter quantity, which has been studied on etched surfaces of germanium,[27] gives one fairly direct information about the sign of the surface potential. When one shines light on a piece of semiconductor, the surface charges negatively for n-type material and positively for p-type. The photo-voltage seen by the grid surrounding the sample in our experiment, consisted of two parts, one due to the change in contact potential with light (ΔCP) and one due to the photo conductivity of the sample. If we apply a voltage to the sample, then the measured signal on the grid is given by $I\Delta R_{\text{Photo}} \pm \Delta CP$ assuming a constant current I through the sample, where the sign of the second term depends on whether the applied sample voltage was positive or negative and on the majority carrier type in the surface region. Hence, by measuring the voltage on the grid with the sample biassed alternately positive and negative, we obtain both the sign and the magnitude of ΔCP. The magnitude was always checked with zero bias on the sample and was found to agree within experimental error with the average of the two measurements made with an applied voltage.

Using the nomenclature of BRATTAIN and GARRETT[27] we have calculated values of δ, proportional to the number of injected carriers from $\delta = wt\rho_i\Delta V_2/Il_{\text{III}}\rho_0^2$ where w is sample width, t the thickness, ρ_0 the resistivity, I the d.c. flowing through it, l_{III} the illuminated length, ρ_i the intrinsic resistivity and ΔV_2 the mean photoconductivity signal voltage. Thus in the absence of traps δ^{-1} is proportional to the surface recombination velocity. From ΔCP and the added carrier parameter δ we have calculated the ratio $\Delta CP/\delta$ (in units of kt/e). The observed ΔCP also includes a Dember potential[28] which occurs between the illuminated and non-illuminated parts of the body of the semiconductor, given by $(b-1)/(\lambda+b\lambda^{-1})$ where b is the ratio of electron to hole mobility and $\lambda = p_0/n_1$. For high resistivity silicon this correction is of the order of a millivolt and has been neglected in the present work. Hence, the values of $\Delta CP/\delta$ are equivalent to $dY/d\delta$. In Table 3, we have listed values of the quantities δ, σ and $dY/d\delta$ obtained before and after various bombardment treatments with 1000 V argon ions. In all cases the pre-bombardment treatment consisted of one minute heating at 1580°K so that the thermally produced bulk donors were present and the bulk resistivity was 280 Ω-cm ($\lambda = 825$, $U_B = -6{\cdot}7$). U_B is given by $U_B = (E_F - E_i)/kT$.

The values of δ (which is a measure of the surface recombination velocity) obtained before and after bombardment were essentially identical. This does not necessarily mean that s was unchanged by the bombardment since the trap density was sufficiently high that it would have obscured changes in s. As described above in Section (c), this is the principal disadvantage associated with the thermally produced donors.

The values of $dY/d\delta$ were negative after bombardment. The negative sign tells us that the surface is becoming more n-type. After bombardment a zero bias photo signal was always detected which indicated the presence of a p–n junction in

Table 3

Run	Ion current (μA)	Time (min)	N_I (ions/cm^2)	δ	$\Delta(dV/d\delta)$ (kT/e)	$\Delta\sigma$ (μmho/square)
	15	0·25	$2{\cdot}2\times10^{14}$	1·2	−0·086	−25
	90	0·25	$1{\cdot}4\times10^{15}$	1·2	−0·091	−30
A	40	2·0	$4{\cdot}8\times10^{15}$	1·5	−0·082	−29
	90	5·0	$2{\cdot}9\times10^{16}$	1·7	−0·081	−30
	900	10·0	$5{\cdot}8\times10^{17}$	1·5	−0·095	−36
	100	5	$3{\cdot}0\times10^{16}$	1·3	−0·089	−22
B	80	5	$5{\cdot}5\times10^{16}$	1·3	−0·084	−27
	60	10	$9{\cdot}0\times10^{16}$	1·6	−0·090	−27
C	1400	90	$7{\cdot}6\times10^{18}$	0·9	0·090	−36
	1000	60	$1{\cdot}12\times10^{19}$	0·9	0·110	−37
	5	0·083	$2{\cdot}5\times10^{13}$	0·4	−0·070	−16
	8	0·166	$1{\cdot}1\times10^{14}$	0·4	−0·14	−23
	20	0·333	$5{\cdot}1\times10^{14}$	0·4	−0·11	−20
D	20	0·500	$1{\cdot}1\times10^{15}$	0·5	−0·12	−22
	50	1·66	$6{\cdot}1\times10^{15}$	0·5	−0·10	−25
	30	10·0	$2{\cdot}4\times10^{16}$	0·5	−0·11	−23
	1000	10·0	$6{\cdot}2\times10^{17}$	0·5	−0·12	−22
	1000	60·0	$4{\cdot}4\times10^{18}$	0·6	−0·11	−24

A: Sample heated after each bombardment; *B*, *C* and *D*: sample heated only before first bombardment. In all four runs the same sample was used.

the surface region. In a study of the rectification properties of bombarded silicon surfaces, OHL[26] found that the back voltage increased steadily until a total charge of 50 μC cm^{-2} had reached the surface. This corresponds to about 3×10^{14} ions cm^{-2}.

If ion bombardment leads to an *n*-type surface then we can deduce the approximate initial condition of the surface from the change in conductivity. Fig. 7 shows the calculated variation in surface conductivity (in mho/square) as a function of the surface potential Y obtained by an extrapolation of the data of KINGSTON and NEUSTADTER[29] using bulk carrier mobilities. The average observed change in surface conductance was 30 μmho/square and comparing this with Fig. 7 we see that before bombardment the surface potential Y must have been about −9 (or 0·24 eV) or in other words the valence band edge at the surface was about 0·13 eV *below* the Fermi level. This result should be compared to that obtained by DILLON and FARNSWORTH[1] who, on a bombarded and annealed surface, found a work function of 4·73 eV and a photoelectric threshold of about 4·6 eV. Since the work function measures the distance from the Fermi level to the vacuum level while the photoelectric threshold measures from the valence band edge to the vacuum level, this implies that the valence band edge at the surface was about 0·13 eV *above* the Fermi level. It is possible that the difference lies in the elimination of the boron doped layer in the present work but since our data are preliminary and at best semiquantitative, no further discussion will be attempted.

(e) *Effect of atomic hydrogen on surface electrical property*

A number of workers have studied the adsorption of atomic and molecular hydrogen[3, 30, 31] on silicon and germanium and the resulting changes in surface conductivity,[32] photoelectric threshold[33] and work function.[1] For both silicon and germanium, molecular hydrogen is adsorbed to only a small fraction of a monolayer if at all. On

VOL. 14 1960

the other hand if a hot filament is present, sufficient numbers of hydrogen atoms are formed that a monolayer can readily be obtained.

Surface conductivity studies on germanium[32] showed that atomic hydrogen acted as an acceptor in that it increased the p-type conductivity of a

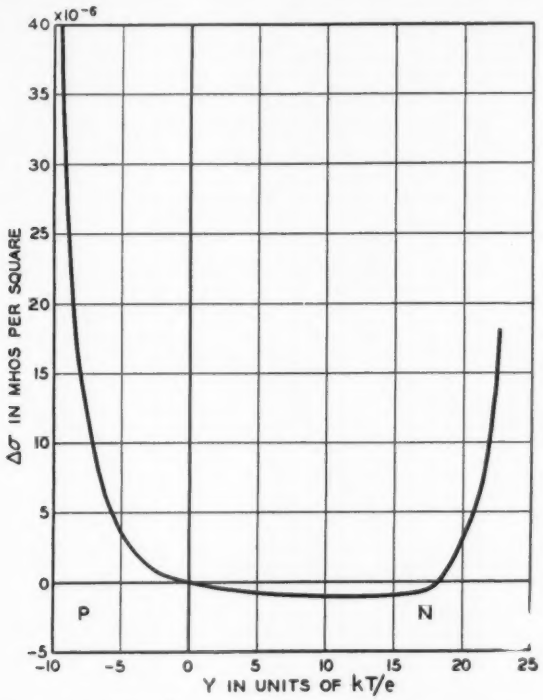


FIG. 7. The calculated change in surface conductivity as a function of surface potential Y, for 280 Ω-cm p-type silicon. When $Y = 0$, $U_S = U_B$ and the energy bands are flat right out to the solid-vacuum interface.

bombardment cleaned surface. On a surface which had been cleaned by heating alone the effect was reversed and the conductivity decreased. It is worth noting that to obtain the largest effect after bombardment, it was necessary to anneal the surface at 650°C for half an hour and at 600°C for 3–5 hr. The results so far obtained on silicon appear to be quite different. DILLON and FARNSWORTH[1] reported that atomic hydrogen decreased the work function from 4·7–4·85 eV to 4·65–4·80 eV. EISINGER measured the photoelectric threshold and found an increase from 4·7 eV to 5·1 eV. Taking these two sets of data together we find that the adsorption of atomic hydrogen has pushed the valence band edge at the surface some 0·4 eV below the Fermi level, assuming that they initially coincided, and that the external barrier has increased by roughly the same amount. These numbers may be in error by 0·1 eV but in any case the change in Y caused by atomic hydrogen is from $+12$ to $+15$ (in units of kT/e).

We have measured surface conductivity changes produced by atomic hydrogen on samples cleaned both by heating and by positive ion bombardment. In the latter case values of ΔCP have also been obtained. During the conductivity measurement on the heated surfaces, the number of hydrogen atoms adsorbed was measured by flashing the sample in a known volume V when the pressure rise in the system ΔP, is related to the number of atoms desorbed by the flash, ΔN_A, by[34]

$$\Delta N_A = 6{\cdot}4 \times 10^{19}\, V \Delta P / A$$

where A is the area of the filament. Hence one can construct a curve of $\Delta\sigma$ against N_A. In Fig. 8 such results which were obtained on a thermally cleaned 10 Ω-cm p-type sample are shown. For reasons previously described,[5] the surface of this material was probably very strongly p-type with threshold increases. The conductivity changes are very large (~ 240 μmho/square) and would require shifts in Y of at least $+10$ which, however, is still within the range predicted by the work function and threshold data. At present we have no explanation for the insensitivity of the electrical properties

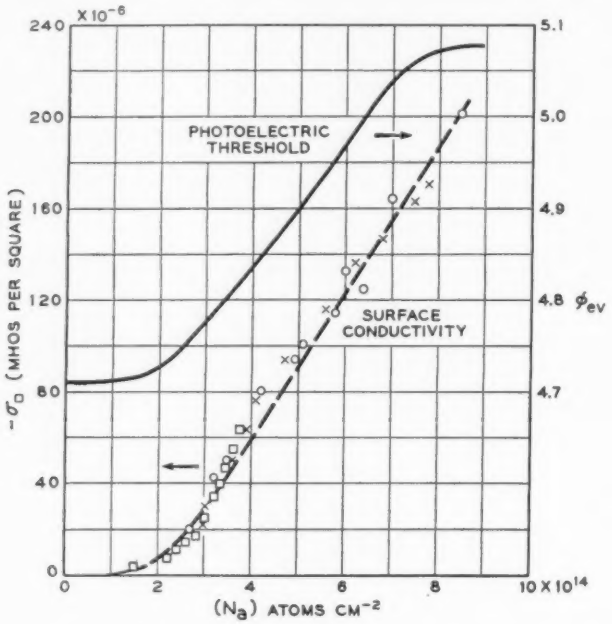


FIG. 8. The change in surface conductivity and photoelectric threshold as a function of the number of adsorbed hydrogen atoms at 300°K. All data were taken on heated samples with strong chemical p-films.

a resistivity much less than 1 Ω-cm. These conductivity changes unlike the germanium results[32] are all negative but are in qualitative agreement with the work function and photoelectric threshold data. Also shown in Fig. 8 is the latter quantity as a function of surface coverage.[33] Since we know very little about the relative number of chemical impurities in the surface regions of the various samples used, an exact comparison between the two sets of data is not possible. They have roughly the same form in that neither changes until a coverage of about 2×10^{14} atoms cm^{-2} is reached, after which the conductivity decreases while the to the adsorption of the first 2×10^{14} atoms cm^{-2}.

In the case of a bombarded and annealed surface on a 280 Ω-cm p-type sample where the highly doped p-layer is absent, the Fermi level is probably 0·1–0·2 eV above the valence band edge at the surface rather than coinciding with it. The adsorption of atomic hydrogen on such a surface also caused the surface conductivity to decrease but by only 38 μmho/square after the adsorption of a monolayer. During the adsorption, ΔCP became steadily more negative, again indicating a downward movement of the energy bands at the surface.

VOL.
14
1960

REFERENCES

1. DILLON J. A. and FARNSWORTH H. E., *J. Appl. Phys.* **29,** 1195 (1958).
2. FARNSWORTH H. E., SCHLIER R. E., GEORGE T. H. and BURGER R. M., *J. Appl. Phys.* **29,** 1150 (1958).
3. LAW J. T. and FRANCOIS E. E., *J. Phys. Chem.* **60,** 353 (1956).
4. ALLEN F. G., EISINGER J., HAGSTRUM H. D. and LAW J. T., *J. Appl. Phys.* **30,** 1563 (1959).
5. ALLEN F. G., BUCK T. M. and LAW J. T., to be published.
6. ALLEN F. G., *J. Phys. Chem. Solids* **8,** 119 (1959).
7. SOULEN J. R., STHAPITANONDA P. and MARGRAVE J. L., *J. Phys. Chem.* **59,** 132 (1955).
8. MARGRAVE J. L., *J. Phys. Chem.* **60,** 715 (1956).
9. LAEGRIED N., WEHNER G. and MECKEL B., *J. Appl. Phys.* **30,** 374 (1959).
10. HORNBECK J. A. and HAYNES J. R., *Phys. Rev.* **97,** 311 (1955).
11. HAYNES J. R. and HORNBECK J. A., *Phys. Rev.* **100,** 606 (1955).
12. FULLER C. S., DITZENBERGER J. A., HANNAY N. B. and BUEHLER F., *Acta Met.* **3,** 97 (1955).
13. SHULMAN R. G., *Semiconductors*, p. 505. Reinhold New York (1959).
14. KAISER W., KECK P. H. and LANGE C. F., *Phys. Rev.* **101,** 1264 (1956).
15. KAISER W., *Phys. Rev.* **105,** 1751 (1957).
16. FULLER C. S. and LOGAN R. A., *J. Appl. Phys.* **28,** 1427 (1957).
17. HONIG R. E., *J. Chem. Phys.* **22,** 1610 (1954).
18. FULLER C. S. and DITZENBERGER J. A., *Phys. Rev.* **91,** 193 (1953).
19. DASH W. C., *J. Appl. Phys.* **27,** 1193 (1956).
20. TWEET A. G., *Phys. Rev.* **106,** 221 (1957).
21. COLLINS C. B. and CARLSON R. O., *Phys. Rev.* **108,** 1409 (1957)
22. WOODBURY H. H. and TYLER W. W., *Phys. Rev.* **105,** 84 (1957).
23. CUSSINS W. D., *Proc. Phys. Soc., Lond.* **68,** 213 (1955).
24. GIANOLA U. F., *J. Appl. Phys.* **28,** 868 (1957).
25. LAW J. T. and GARRETT C. G. B., *J. Appl. Phys.* **21,** 656 (1956).
26. OHL R. S., *Bell Syst. Tech. J.* **31,** 104 (1952).
27. BRATTAIN W. H. and GARRETT C. G. B., *Bell Syst. Tech. J.* **35,** 1019 (1956).
28. ROOSBROECK W. VAN, *Phys. Rev.* **101,** 1713 (1956).
29. KINGSTON R. H. and NEUSTADTER S. F., *J. Appl. Phys.* **26,** 718 (1955).
30. LAW J. T., *J. Chem. Phys.* **30,** 1568 (1959).
31. LAW J. T., *J. Phys. Chem.* **59,** 543 (1955).
32. HEILAND G. and HANDLER P., *J. Appl. Phys.* **30,** 446 (1959).
33. EISINGER J., *J. Chem. Phys.* **30,** 927 (1959).
34. BECKER J. A., *Advanc. Catalys.* **7,** 162 (1955).

VOL
14
196

VOL.
14
1960

DISCUSSION

R. FORMAN (Union Carbide Corporation): Can the data shown in Fig. 2 of the paper be explained by bombardment damage? The treatment described, using a high bombardment rate and time, is a drastic one. It could be capable of causing deep mechanical damage in the sputtering process. If this were so, could one explain the rise in resistivity as a phenomenon associated with the development of a badly damaged surface layer with a higher resistivity. If this were so, then curve (2) in Fig. 2 would imply that the damage done to the surface by the annealing cycle has not been eliminated. We suggest this as a possible explanation on the basis of data we have obtained on argon bombardment treated germanium in which the sample was exposed to a much milder sputtering treatment. As a result of this mild treatment the effective surface mobility was drastically reduced which we have interpreted as due to mechanical damage caused by the argon bombardment.

REPLY: The data of Fig. 2 at the high ion density end cannot be due to bombardment damage, but the small change for ion densities of less than 10^{15} cm^{-2} could be caused by either damage or the embedding of argon ions in the surface layer. The argument for the right hand part of the curve goes as follows. We know that after heating silicon in vacuum a strongly doped boron layer is produced on the surface and the sample resistance will drop in a typical case from 3×10^5 to about 1×10^5 Ω. After either chemically etching or heavy ion bombardment the value returns to approximately 3×10^5 Ω and will stay in this region during a series of bombardments and anneals. If one were producing a badly damaged layer by bombardment the sample resistance would be expected to rise above its initial value, a situation which does not arise under suitably controlled conditions. In addition, the correlation between the resistance changes caused by etching and bombarding a heated surface would not exist.

It is true that bombardment will cause changes in conductivity but unless the depth of mechanical damage approaches several microns one will never obtain conductivity changes on high resistivity silicon greater than several tens of μmho/square. For the ion energies used a reasonable damage depth is only 10^{-6} cm (Table 2).

J. A. DILLON, Jr. and H. E. FARNSWORTH (Brown University): In the past few years ion-bombardment cleaning has been applied to a number of metal and semiconductor crystals and the results monitored by electron-diffraction techniques. A consideration of the results of these experiments suggests several important points.

(1) For most of these materials it has been possible to determine a set of operating conditions under which bombardment and annealing can be carried out without introducing an important number of surface defects.

(2) These conditions are not the same for all materials (e.g. the annealing temperature for silicon is much higher than that for germanium), and it is well known that under certain other conditions ion bombardment can introduce surface damage which cannot be annealed out. For example, silver and gold are known to be very malleable, and the lattice structures are easily damaged. Hence, one may expect recrystallization or permanent damage under certain conditions of ion bombardment.

(3) It is therefore imperative that the bombardment cleaning be performed only under those conditions which are known to result in clean, defect-free surfaces, and it should not be at all surprising if satisfactory results are obtained under other conditions. This is particularly true when the technique is applied to new materials. Objective evidence should be obtained before a combination of bombardment and annealing parameters is assumed to be satisfactory.

(4) Beyond the question of the possible introduction of defects, it should also be kept in mind that certain surface electrical properties (e.g. surface recombination velocity) are more sensitive to annealing conditions than others and that the effects of heat treatment are more complex for some materials than for others.

Therefore for each material the bombardment-cleaning conditions should be carefully investigated, and extrapolating the results from one material to another should be considered as having dubious value.

REPLY: We are in complete agreement with the comments made, and they state precisely what we have been attempting to do. That is, by making electrical measurements sensitive to the surface condition, define the bombardment and annealing conditions necessary to obtain a sample with three-dimensional homogeneity and a clean surface. We feel that in so far as the former is concerned (especially with respect to such things as chemically doped layers) electrical measurements are much more sensitive than electron-diffraction when it comes to detecting impurities in the 10^{12}–10^{13} cm^{-2} range.

J. A. DILLON, Jr. and R. M. OMAN (Brown University): In studies of the photoelectric properties of silicon surfaces* we have reported evidence for the existence of diffused layers due to high-temperature heating and advised that crystals be cleaned by ion bombardment *prior* to such heating. These results are of course consistent with the detailed observations reported by Law. More recent experiments in our laboratory have also been directed toward effects on the conductivity of silicon surfaces induced by heating in vacuum. Pertinent to the present discussion is the fact that if the argon ion bombardment (500 eV, 100μA/cm^2) is performed *before* heating at 1000°C, the number of incident ions/cm^2 required to prevent the formation of high-conductivity films on subsequent heating is at most 10^{18}/cm^2. It should be noted that Law's results which indicate a

* DILLON J. A., Jr. and FARNSWORTH H. E., *J. Appl. Phys.* **29,** 1195 (1958).

necessary bombardment of at least $10^{19}/cm^2$ apply to a sample which had been heated before bombardment. Once the surface contaminant has diffused into the silicon the whole diffused layer must be sputtered off, whereas prior to heating only the contaminant must be removed. This point is emphasized lest some of the previous experiments, which have actually satisfied conditions necessary for film removal, be too hastily discounted.

REPLY: We have made only preliminary attempts to remove the surface contamination after bakeout, but before heating. With ion densities near 10^{17} ions/cm^2 the conducting p-type film was still formed on subsequently heating the sample to 1500°K. We should like to emphasize that some sensitive measurement (such as sample conductance on p-type material) needs to be used to ensure that added acceptors are not present in the surface region and if one is cleaning by ion bombardment that the opportunity exists for measuring the equivalent of a concentration profile of electrically active impurities.

A. KOBAYASHI (Electrical Communication Laboratory, Tokyo, Japan): Dr. AHEARN* has reported that he could separate the surface contamination from the bulk impurity. I would like to know if it be able to detect the contamination of the surface cleaned by heating.

REPLY: We have examined several heated samples by this technique and have found weak boron and also aluminum lines from the surface region when these lines were missing from both the bulk and from controls. The major problem is sensitivity as we are dealing with considerably less than a monolayer of contamination.

P. HANDLER (University of Illinois): It has been observed by workers at the University of Illinois that rapid cooling of germanium from 600°C to room temperatures will radically alter the bulk resistivity. In order to obtain reproducible results it was found that the sample had to be cooled slowly over a period of at least one half-hour between 600°C and 300°C. It would be interesting to know whether the author has ever observed the effects of slow cooling on his samples.

REPLY: The effect on the sample resistivity of slow versus fast cooling, was observed in early measurements on silicon samples. It was found, however, that most of this effect could be eliminated if the sample were rigidly supported at only one end. Hence it was presumably due to dislocations introduced by the sample bending during the heating cycle. The residual thermally caused changes above 1000°K are discussed in the text where it can be seen that slow cooling in the 1580–1000°K range are very important.

---

* AHEARN A. J., *A.V.S. 6th National Vacuum Symposium at Philadelphia*, 1959.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 23–26. Printed in Great Britain.

SURFACE MEASUREMENTS ON FRESHLY CLEAVED SILICON *p–n* JUNCTIONS

G. W. GOBELI and F. G. ALLEN

Bell Telephone Laboratories, Murray Hill, New Jersey

Abstract—Silicon (111) faces 1·5 mm × 6 mm in area and free of visible imperfections, containing a *p–n* junction, have been produced by cleavage in high vacuum. The cleavage features and the technique are discussed. Contact potential difference (Kelvin method) across the junction indicates bending of the bands at the surface due to surface states by many tenths of a volt.

VOL. 14 1960

1. INTRODUCTION

A LARGE segment of recent semiconductor surface work has centered on the production and study of the atomically clean surface. There are three current methods for the production of a clean silicon surface: (1) lengthy positive ion bombardment followed by degassing and annealing at moderate temperatures for long times[1, 2] (2) vacuum heating to temperatures just below the melting point,[3] (3) cleavage in high vacuum.[4] The first two methods suffer mainly from heat treatment effects and possible in or out diffusion of impurities which alter the crystal properties, particularly at the surface. Thus, it is possible to produce a crystal which is atomically clean in the two-dimensional surface but which is contaminated in the third dimension perpendicular to that surface.

The cleavage technique eliminates this problem as well as the problem of demonstrating that the surface is initially atomically clean. It, however, suffers from three difficulties: (1) the cleaved areas are usually quite small and imperfect and sample geometries are not well suited to electrical measurements, (2) the surface cannot be regenerated at will save by another cleavage and (3) possible existence of physical damage to the crystal structure such as plastic deformation, dislocation production, fracture cracks, etc. These three problems are discussed in some detail in this work.

2. CLEAVAGE OF SILICON

(a) *Discussion of the cleavage method*

Both silicon and germanium, like diamond, cleave only in the (111) plane. The tendency to cleave rather than fracture is very slight so that great care must be exercised to produce consistently good cleaved surfaces.

The usual technique of cleaving by driving a wedge shaped knife into a notch cut into the sample was discarded since it almost never produced a cleaved surface of the quality thought necessary for the measurements to be made. The method which was employed and which ultimately proved reliable in yielding good usable surfaces was that of clamping a long (111) oriented silicon specimen and fracturing it by application of a bending force. Samples so fractured exhibited the true flat cleavage surface only in that area of the cross section which was under tension just before the cleavage crack started through the crystal, i.e. the upper half of the newly exposed cross section. The lower half, which was under compression exhibits extremely rough fracture characteristics. This fact is most strikingly demonstrated when it is observed that a sample exhibits the flat cleavage characteristic over its total cross section when it is exposed by a purely tension fracture, i.e. purely axial force until the sample fails. Such a cleavage is difficult to produce in vacuum, however.

As has been reported by others[5] the condition of the tension surface of the sample is extremely important in so far as it affects the cleavage perfection of the exposed surface. The major imperfections in the flat portion of the exposed face are tear marks, i.e. very large fracture steps, radiating from a point in the tension surface of a sample fractured by bending. These marks, radiating

from a single point, indicate that the cleavage initiates from that point when the local tension exceeds a critical value in the neighborhood of an imperfection in the surface. This assumption is somewhat supported by the observation that if a moderate bending force is applied to a carefully etched sample, the sample can remain in the stressed condition for a considerable time if left undisturbed. However, the cleavage may be spontaneously initiated at any single given point by simply touching the tension surface lightly with a pointed object and the tear marks originate from the point at which the surface was touched. Conversely a lapped or sandblasted sample fractures at much lower forces than etched samples and in general the tear marks originate from several points in the tension surface.

These considerations led to the use of samples with the cross section shown in Fig. 1. The sample

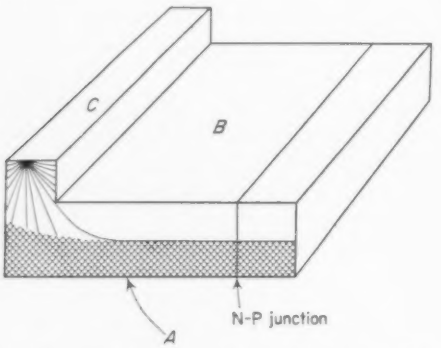


FIG. 1. Diagram of cleavage sample geometry showing overall features of cleaved surface.

is clamped with the clamps resting firmly and evenly on the surfaces *A*, *B* and *C*. A small scratch is scribed on the surface *C*, then a sudden force is applied on surface *B*, 1 cm beyond the supporting edge, so that *B* and *C* become the tension surfaces. The greatest tension occurs in surface *C* and fracture initiates at the scratch. The tear marks radiating from the fracture point are blocked from the portion of the sample below surface *B* and if that surface is very carefully prepared, the clear area shown in the cross section of Fig. 1 is free of gross cleavage imperfections.

(b) *Examination of the cleaved surfaces*

The flat portions of the fractured surfaces were examined under a metallographic microscope at magnifications varying from 10× to 1000×. Except for an occasional large cleavage step found in some of the samples examined, the areas were so featureless that focussing could only be accomplished at edges or on dust particles, even when glancing illumination was used.

Figure 2 shows four photographs of the cleaved surfaces taken with an optical microscope illustrating features of the cleavage.

Reflection electron diffraction pictures were taken along the length of the flat cleaved area. Large numbers of well-developed Kikuchi lines were observed as well as sharp spots corresponding to the (111) silicon spacing. This is indicative of a high degree of perfection in the last few atomic layers near the surface. Especially interesting is the complete lack of any evidence for amorphous or foreign material in or on the surface.

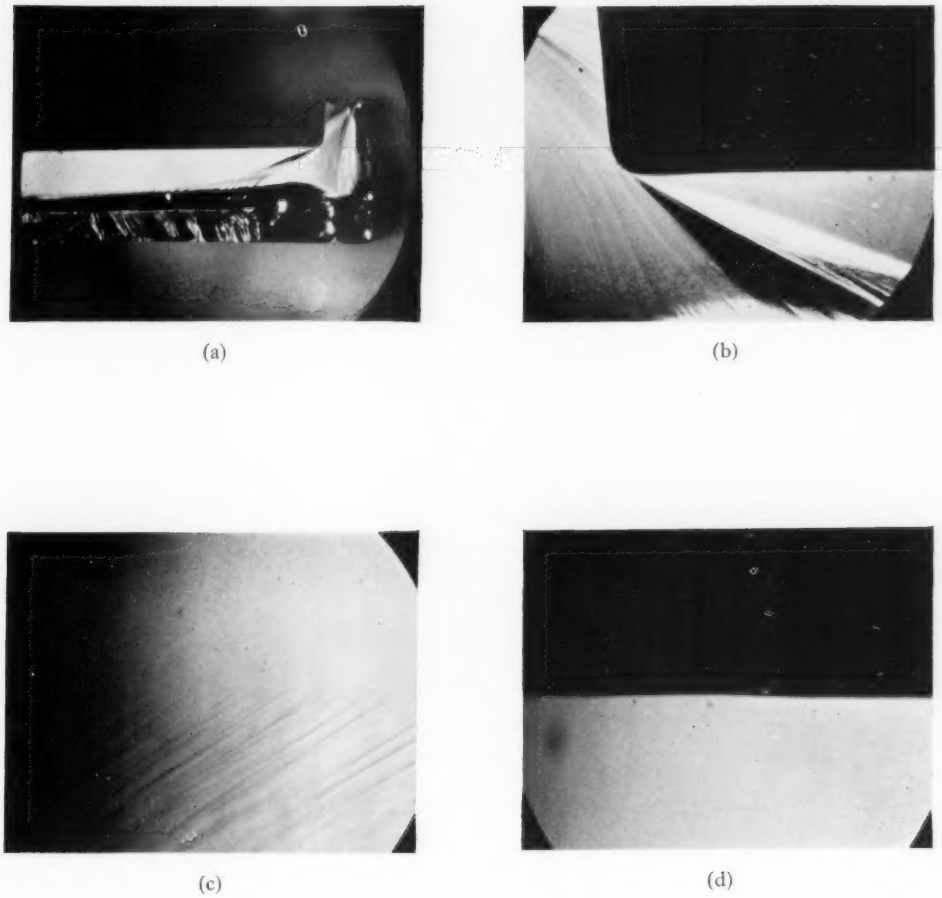
Electron micrographs were taken of platinum-shadowed evaporated-carbon replicas. They show that selected areas of the surfaces are flat within the resolution of the grainy structure of the replica. An imperfection or step of about 50 Å amplitude would be detectable. Large numbers of cleavage steps varying in height from 50 Å to 500 Å separated by distances of thousands of angstroms are observed in all surfaces with the areas between steps being flat within the limits described above. Fig. 3 (a) and (b) taken at 6000× illustrate the step structure, while Fig. 3 (c) at 62,000× illustrates a flat area between steps. Only a few cracks were observed in one sample of three examined and one of these is shown in Fig. 3 (d). There was no evidence whatever for plastic deformation.

Therefore, although the surfaces are not as featureless on the submicroscopic level as might be hoped, there is no evidence for physical damage to the crystal structure resulting from the fracture of the sample.

3. EXPERIMENTAL PROCEDURE

The sample was mounted on the end of a rod which could be manipulated through a sylphon bellows; a combination tungsten carbide scribe and blunt chisel tool operated through a second bellows was used to scribe and fracture the sample. These operations were carried out in a bakeable vacuum system in which residual

VOL.
14
1960

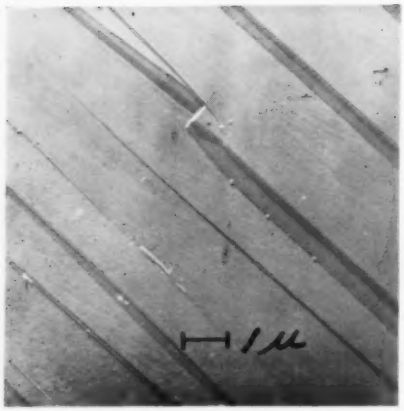


(a) (b)

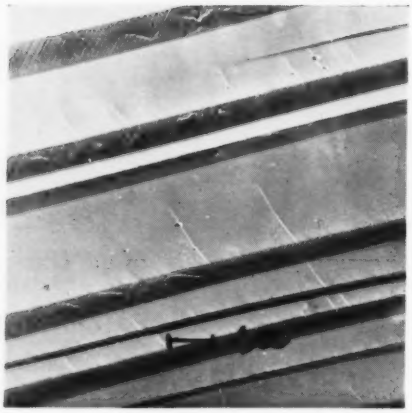
(c) (d)

FIG. 2. Optical microscope photographs of cleaved silicon surface: (a) overall view 10×; (b) 120× surface B near corner to surface C; (c) 1000× region at tension and compression boundary; (d) 1000× region including surface C.

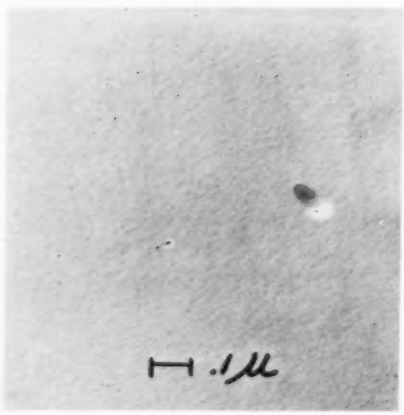
[*facing p.* 24



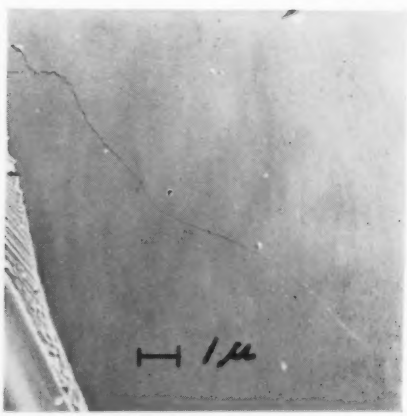
(a)



(b)



(c)



(d)

FIG. 3. Electron micrographs of cleaved silicon surfaces. Shadowing angle is 63·5°. Grainy background is believed to be inherent in replica and structure in Fig. 3 (c) is an artifact used for focussing.

VOL. 14 1960

pressures of $1{\cdot}5\times10^{-10}$ mm Hg were normally maintained. The pressure, as measured by a Bayard–Alpert type ionization gauge, was recorded continuously during the fracture procedure. Upon cleavage pressure bursts of 1×10^{-9} mm Hg were typically obtained and pump down to the initial vacuum occurred in a few seconds. Similar pressure bursts were often observed during scribing and other manipulations in the tube.

The cleaved surface cuts across a p–n junction of arbitrary doping so that both p- and n-type surfaces are simultaneously exposed for comparison. Contact potential is measured by using a small hemispherical molybdenum boss as a reference surface vibrated close to the sample surface. This microprobe technique is used in order to be free of spurious effects from the contaminated outer surface of the sample although at present it limits the sensitivity of the measurement to about $\pm0{\cdot}010$ V.

As indicated in Fig. 4 the contact potential difference across the p–n junction gives the difference in work function, $\Delta\phi_{NP}$, of the p- and n-sides directly. The contribution of the *surface dipole*, shown as χ in Fig. 4, remains constant on both sides of the junction and the bands are arbitrarily shown bending to produce a p-type surface. This work function difference, in comparison with the known Fermi level positions relative to the band edge on either side of the junction in the bulk give some information concerning the surface states found on a clean (111) silicon face.

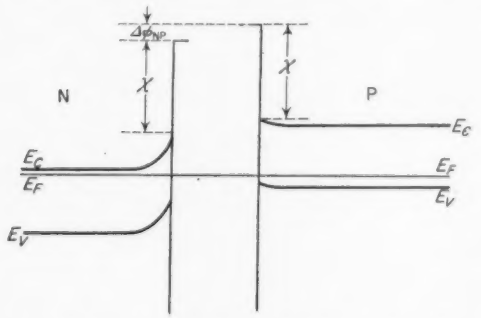


FIG. 4. Energy level diagram for both sides of p–n junction.

Results thus far on several samples from each of two heavily doped junctions show definite and reproducible work function differences from p- to n-sides in the expected direction, that is $\phi_P > \phi_N$, but only 1/5 to 1/10 of the differences in E_F-E_V in the bulk. The samples, dopings and differences are shown in Table 1. The work function difference is larger for the more heavily doped junctions as expected.

The threshold for external photoelectric emission was measured on the hemispherical molybdenum reference surface. This surface was polycrystalline and uncleaned, and due to beam defocussing difficulties it cannot be regarded as reliably characteristic of the tip of the molybdenum boss reference. The measurements indicated a photo-threshold of about 4·6 V and from the contact potential difference value we find the work function of the heavily doped (0·007 Ω-cm) p-type (111) face, in the freshly-cleaved condition, to be 4·7 V. This is in rough agreement with previous results of ALLEN *et al.*[3, 6] for a silicon surface cleaned by heating alone in high vacuum and known to have a strong p-film at the surface. Our strongly n-type (0·04 Ω-cm) (111) surface, freshly cleaved, had a work function 0·15 V lower than the p-type side.

The work function values of all silicon cleaved surfaces drifted upward 0·1–0·2 V after 3–5 hr exposure to the residual gases at a pressure of

Table 1

		Difference in E_F-E_V across junction in bulk	Difference in work function, freshly cleaved $\phi_P-\phi_N$
Sample 1	0·007 Ω-cm p-type 0·04 Ω-cm n-type	1·0 V	0·15 $\pm$0·03 V
Sample 2	0·42 Ω-cm p-type 4·5 Ω-cm n-type	0·7 V	0·10 $\pm$0·03 V

3×10^{-10} mm Hg. Upon exposure to oxygen at a pressure of $\sim 10^{-6}$ mm Hg for several minutes, the silicon work function increased again by 0·3–0·4 V, giving a total change in work function of 0·5–0·6 V—again in agreement with other workers' findings on surfaces cleaned by sputtering and heating. The difference in work function between *n*- and *p*-sides tended to decrease somewhat on standing in vacuum, and to disappear altogether upon exposure to oxygen. This effect was more pronounced in the more lightly doped sample than in the heavily doped one.

The difference in contact potential measured on the more heavily doped junction requires the assumption of $\sim 2 \times 10^{12}/\text{cm}^2$ surface states below the Fermi level to produce the required bending of the bands.

These measurements in combination with experiments on the threshold for external photoelectric emission which are presently in preparation should enable one to locate accurately the position and density of the surface states present on a freshly cleaved (111) silicon surface.

In conclusion it may be said that for freshly-cleaved silicon surfaces in high vacuum the differences in work function between *n*- and *p*-type samples are measurable and in the expected direction. They are large enough (0·15 V) to show that clean *n*-type samples do not have degenerate *p*-type surfaces. They are not nearly so large, however, as would be predicted from a naive picture assuming no surface states and a flat-band model at each surface. Our data indicate bending of the bands at the surface due to surface states by many tenths of a volt.

Acknowledgements—The authors wish to thank C. J. Calbick, Mrs. M. Read and Miss S. E. Koonce for examinations with the electron microscope; and G. J. Wolff and A. A. Studna for their technical assistance.

REFERENCES

1. Farnsworth H. E., Schlier R. E., George T. H. and Barger R. M., *J. Appl. Phys.* **29,** 1150 (1958).
2. Law J. T., this Conference, p. 9.
3. Allen F. G., Eisinger J., Hagstrum H. D. and Law J. T., *J. Appl. Phys.* **30,** 1563 (1959).
4. Simon R., *Bull. Amer. Phys. Soc.* ser. 2 **4,** 410 (1959).
5. Johnston T. J., Stokes R. J. and Li C. H., *Acta Met.* **6,** 713 (1958).
6. Allen F. G., *J. Phys. Chem. Solids* **8,** 119 (1959).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 27–32. Printed in Great Britain.

ELECTRICAL PROPERTIES OF CLEAVED GERMANIUM SURFACES*

D. R. PALMER, S. R. MORRISON and C. E. DAUENBAUGH

Honeywell Research Center, Hopkins, Minnesota

Abstract—The effects of oxygen on the electrical properties of a cleaved germanium surface have been studied by measuring the field effect, conductance, and photoconductance. Upon cleavage, the surface is more *p*-type than the bulk and remains unchanged in a vacuum of about 10^{-9} mm Hg or less. When initially exposed to oxygen, the conductance tends toward a more *p*-type surface; then the surface changes towards *n*-type continuously at a constant oxygen pressure. When the system is evacuated, the surface reversibly becomes more *p*-type, indicating that oxygen tends to make the surface more *n*-type. In addition, the slow relaxation of the field effect in oxygen almost disappears in vacuum. The surface eventually becomes *n*-type in oxygen at 15 mm Hg for all of the samples measured, an effect which has not apparently been observed on surfaces cleaned by argon bombardment. The results are interpreted in terms of acceptor states at the surface and donor states which depend on sorbed oxygen and produce a slow relaxation of the field effect.

VOL. 14 1960

INTRODUCTION

THE methods which have been used to obtain "clean" germanium surfaces are ion bombardment,[1, 2] cleavage,[3, 4] and hydrogen reduction.[5] Of these methods, ion bombardment has been the most frequently used means to produce a clean surface. We have chosen to use the cleavage technique because there can be no doubt that the surface is clean immediately after cleavage, although there will be some adsorption taking place even in ultra-high vacuum. It is not clear that we have been able to take advantage of this initially clean surface as yet, for during the cleavage gas is evolved from the moving parts, the pressure of this gas momentarily rising as high as 10^{-7} mm Hg.

Another possible advantage of the cleavage technique over the ion-bombardment technique is the possibility of less surface damage. The only observed dislocations on a cleaved germanium surface as measured by etch pit count[6] are those inherent in the bulk. The damage under ion bombardment has not been clearly defined by workers in the field. Before the annealing procedure, the damage is apparently extensive, but the final damage subsequent to annealing is not known.

* This research was supported in part by the United States Air Force through the Air Force Office of Scientific Research (ARDC).

The present investigations have a twofold purpose. First a comparison of the results to the results with ion-bombarded surfaces is of interest; second, the initial steps in an attempt at quantitative determination of the influence of oxygen on a clean surface are described.

EXPERIMENTAL PROCEDURE

Reasonably good cleaved surfaces are obtained using an arrangement similar to that shown in Fig. 1 which does not include the vacuum envelope. The germanium is oriented and cut so that the desired plane of cleavage is along a (111) plane, a plane of easy cleavage. Notches are cut at the top and bottom; the top notch has a peak in the center and the bottom notch is flat. To cleave the germanium, the hammer is raised with an external electromagnet and dropped. The top chisel has a much smaller angle than the bottom so that cleavage is initiated at the top, presumably at one point.

Upon cleavage, the two parts separate and become a field plate and a sample with a separation of about 0·015 in. This arrangement makes field effect calculations possible since the spacing from field plate to sample should be uniform even if the cleave is not perfectly planar. Leads are attached to the sample at positions indicated by the *x*'s on Fig. 1. This arrangement allows for measurement

of the conductivity of the portion of the sample with a cleaved surface and also a portion of the sample with no cleaved surface. The latter measurement is used to account for resistivity

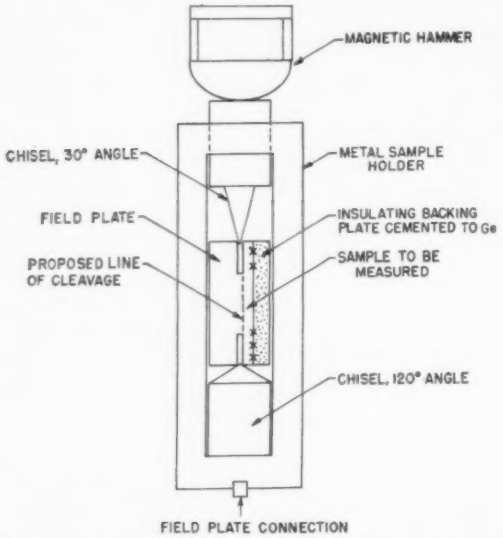


FIG. 1. Cleavage apparatus. X indicates the position of contacts to the sample.

changes due to temperature. All surfaces other than the cleaved surface are lapped or sandblasted so that the former are presumably unaffected by ambient or an applied field. Photoconductance measurements are possible by shining light through a hole in the sample holder. The light is filtered by the germanium field plate so that penetrating light strikes the sample. Preliminary photoconductance measurements have been made in this fashion, but will not be reported here as they have not been adequately confirmed.

We have been using two different systems to investigate the effect of oxygen on clean germanium surfaces. The first, which will be referred to as the "high vacuum" experiment, involves cleaving at a pressure of about 10^{-9} mm Hg and subsequently exposing the surface to oxygen at a pressure less than 10^{-3} mm Hg. The second or "high pressure" experiment involves cleaving at a vacuum of about 10^{-6} mm Hg and subsequently admitting oxygen at 13 mm Hg. Measurements of the field effect and conductivity are made in vacuum and in oxygen. The field effect was measured with a d.c. field so that any slow relaxation of the field effect might be observed.

The insulating backing plate was cemented to the sample to provide additional physical support during cleavage. In the "high pressure" experiment we were able to use a stronger cement and thereby reduce the sample thickness. The dimensions of the samples used in the "high vacuum" experiment were about 6 mm × 3 mm × 2 mm. In the "high pressure" experiment it was possible to obtain samples about 1 mm thick.

EXPERIMENTAL RESULTS

Measurements were made on both 25 Ω-cm p-type and 17 Ω-cm n-type germanium. The results were qualitatively similar for these two resistivities so typical results for the 25 Ω-cm p-type case will be reported here, with the "high vacuum" experiments being considered first.

After cleavage at about 10^{-9} mm Hg, the field effect mobility (μ_{FE}) was about 50–150 cm^2/V-sec p-type and no change in μ_{FE} was observed for at least 100 hr at this pressure. Fig. 2 shows typical results for surface conductivity ($\Delta\sigma_S$) and μ_{FE} as a function of time in oxygen at 10^{-5} mm Hg. The change in surface conductivity was related to an arbitrary zero before oxygen admission. The surface conductivity reached a maximum shortly after exposure to oxygen and then continuously decreased thereafter. Two other experiments yielded peak values for $\Delta\sigma_S$ 12·8 and 12·5 μmho/square. The field effect mobility continuously increased throughout the experiment and reached a final value roughly twice the initial mobility. It should be emphasized here that the absolute value of μ_{FE} in this experiment and those to follow has about a 30 per cent uncertainty because the capacity can not be calculated exactly.

Assayed reagent grade oxygen was admitted through a high vacuum valve after the pressure had been reduced to about 10^{-3} mm Hg. It was necessary to continuously admit oxygen and to continuously monitor the pressure with an ion gauge during the course of our experiments, because the glass walls apparently sorb oxygen in sufficient quantity to drastically lower the pressure in a closed system. We were also able to reduce anomalous ion gauge effects (exclusive of pumping) by introducing a grounded metal baffle between the ion gauge and the sample system. Previously, we

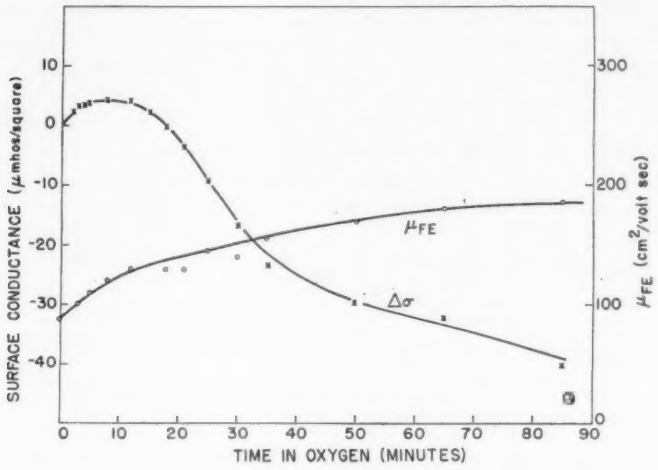


FIG. 2. Surface conductivity and field effect mobility as a function of time in oxygen at 10^{-5} mm Hg. The surface conductivity is related to an arbitrary zero at cleavage.

VOL. 14 1960

had attempted to introduce oxygen into the system by heating a strip of oxidized nickel. We later found that the gas evolved was almost entirely CO_2 and CO, which had little or no effect on the electrical properties of a cleaved germanium surface.

The maximum and subsequent decrease in conductivity observed in these experiments has also been observed by others[7–9] on surfaces cleaned by argon bombardment. The field effect results are similar to the results reported by HANDLER[7]. Thus, qualitatively, the results are similar. However, quantitatively, there seem to be unexplained differences. For example, the maximum conductivity we observed in oxygen relative to vacuum was about an order of magnitude smaller than observed by the Illinois group.[7, 9] This difference in the maximum is quite significant despite the general irreproducibility of the maximum observed on both the argon bombarded and the cleaved surfaces.

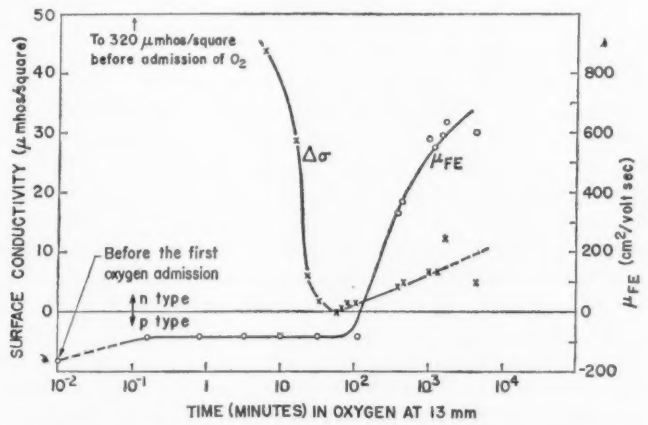


FIG. 3. Surface conductivity and field effect mobility as a function of time in oxygen at 13 mm Hg.

In the "high pressure" experiments the germanium was cleaved at 10^{-6} mm and oxygen was later admitted at 13 mm. Fig. 3 shows $\Delta\sigma_S$ and μ_E as a function of time in oxygen. Here the surface conductivity was related to the minimum conductivity (intrinsic surface) which occurs at about the same time as μ_{FE} changes sign. This experiment indicates the surface was strongly p-type initially and later became n-type in the presence of oxygen. BARNES and BANBURY[10] also observed an n-type surface on cleaved surfaces of p-type germanium.

From the curves shown, it is apparent that there is a reversible and an irreversible reaction occurring at the surface. The reversible phenomenon is relatively rapid and is associated with changing from oxygen to vacuum. The irreversible phenomenon is slower and seems related to the total time oxygen has reacted with the surface.

The observation, consistent with other workers, of an initial p-type surface upon cleavage, indicates a high density of acceptor states on the surface

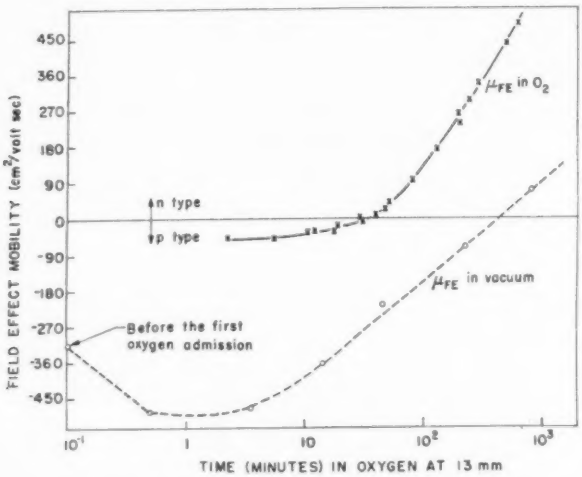


FIG. 4. Field effect mobility as a function of time in oxygen at 13 mm Hg. The solid curve represents readings in oxygen and the dashed curve represents readings in vacuum.

Figure 4 shows μ_{FE} as a function of time in oxygen at 13 mm. In this case the sample was cycled between oxygen and vacuum. The *in vacuum* measurements represent equilibrium vacuum values. These results show that the surface is more n-type in oxygen than in vacuum. Slow relaxation of the field effect was observed when the surface was exposed to oxygen and it was especially noticeable when the surface was n-type. The slow relaxation decreased considerably in vacuum. Qualitatively, similar results were observed on cleaved surfaces of 17 Ω-cm n-type germanium but in this case less time in oxygen was required to reach an intrinsic surface.

which are effective in altering the surface conductivity and field effect mobility. It is apparent from Fig. 3 that the irreversible reaction involves changes in the fast state density. We will make the assumption that the reversible reaction does not involve changes in the density or energy of the dominant fast states. This is supported by the observation that upon oxygen admission after evacuation μ_{FE} returns very quickly to the measured value prior to evacuation.

Assuming that the surface barrier is altered during evacuation without changing the density or energy of the dominant fast states, it should be possible to calculate the density and energy of the dominant fast states from measurements of $\Delta\sigma_S$ and μ_{FE} at any two values during evacuation. A

VOL. 14 1960

preliminary experiment was done on 50 Ω-cm p-type germanium to obtain measurements of this kind. Fig. 5 shows $\Delta\sigma_S$ and μ_{FE} as a function of time in oxygen (measured in O_2). Points A and B represent typical vacuum measurements of $\Delta\sigma_S$ and μ_{FE}, respectively. Surface conductivity versus surface barrier calculations were made using the calculation by KINGSTON and NEUSTADTER[11] for the free carrier concentrations versus surface barrier and the calculations by SCHRIEFFER[12] for 1 hr. It is observed that little variation in surface properties occurred in a reasonable time compared to earlier measurements and with larger changes in $\Delta\sigma_S$ and μ_{FE} better results should become available.

DISCUSSION OF RESULTS

A cleaved germanium surface is initially highly p-type. When exposed to oxygen, the surface initially becomes slightly more p-type followed

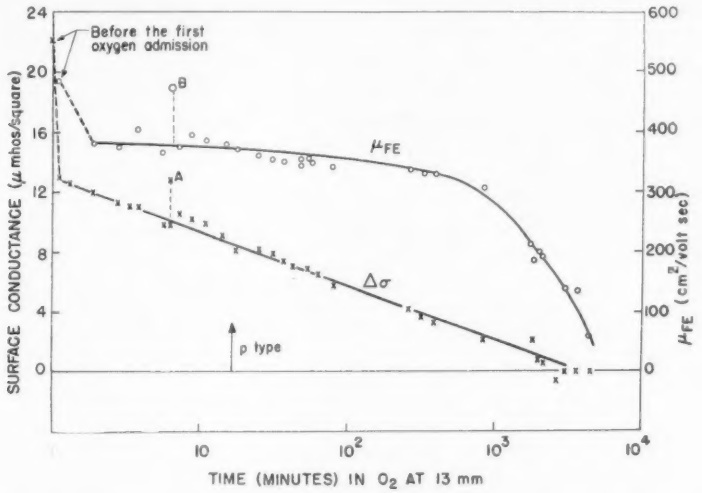


FIG. 5. Surface conductivity and field effect mobility as a function of time in oxygen at 13 mm Hg. Points A and B represent equilibrium measurements in vacuum.

the effective mobility versus surface barrier. The conductivity measurements were then used to determine the surface barrier. Calculations were also made for μ_{FE} as a function of surface barrier, including a correction term to account for the energy and density of surface states assuming only one type of surface states affect the field effect mobility. Corrections for effective mobility were necessary in this case also. Calculations were then made from measurements like those shown in Fig. 5.

The energy of the dominant fast states subject to the assumptions and approximations above was calculated to be $0{\cdot}10 \pm 0{\cdot}02$ eV below the center of the gap. The density decreased from about $3 \times 10^{11}/cm^2$ at 1 min to about one-half this value at

by a continuous change towards an n-type surface. At sufficiently high pressures, a stable n-type surface can be produced which is in contrast to results reported on surfaces cleaned by argon bombardment. It was also observed that a cleaved surface is more n-type in oxygen than in vacuum and that slow relaxation of the field effect, if observed in oxygen, decreases or disappears in vacuum. These results, which are qualitatively similar to those reported on surfaces cleaned by argon bombardment, may be explained, at least in part, by a theory proposed by HANDLER and PORTNOY[9]. They interpreted their results in terms of a high initial density of acceptor type surface states. The initial increase in conductivity was explained by adsorption of oxygen as O^- and

the subsequent decrease in conductivity by donation of an electron to the space charge region and by a decrease in the density of acceptor type surface states. From our results it is necessary to further assume that oxygen at later stages produces some kind of donor states on the surface and that these donor states may be similar to the normal "slow states" observed on etched surfaces.

It is planned to combine the "high vacuum" and "high pressure" experiments in the future so that more control over changes in the surface are possible. It is desirable to extend the calculations of the energy and density of fast states to lower degrees of surface contamination. This may be achieved by combining the "high vacuum" and "high pressure" experiments on the same sample. Our present results concerning the energy and density of the fast states are limited because we cannot make calculations from the measurements immediately after cleavage and large changes occur during the first exposure of the surface to oxygen at 13 mm Hg.

REFERENCES

1. FARNSWORTH H. E., SCHLIER R. E., GEORGE T. H. and BURGER R. M., *J. Appl. Phys.* **26,** 252 (1955).
2. SCHLIER R. E. and FARNSWORTH H. E., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 3. University of Pennsylvania Press, Philadelphia (1957).
3. BARNES G. A. and BANBURY P. C., *Proc. Phys. Soc., Lond.*, **71,** 1020 (1958).
4. PALMER D. R. and DAUENBAUGH C. E., *Bull. Amer. Phys. Soc.* ser. 2, **3,** 138 (1958).
5. KOBAYASHI A. and KAWAJI S., *J. Phys. Soc. Japan* **12,** 1054 (1957).
6. STOKES R. J., private communication.
7. HANDLER P., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 23. University of Pennsylvania Press, Philadelphia (1957).
8. LAW J. T. and GARRETT C. G. B., *J. Appl. Phys.* **27,** 656 (1956).
9. HANDLER P. and PORTNOY W., *Phys. Rev.* **116,** 516 (1959).
10. BARNES G. A. and BANBURY P. C., *J. Phys. Chem. Solids* **8,** 113 (1959).
11. KINGSTON R. H. and NEUSTADTER S. F., *J. Appl. Phys.* **26,** 718 (1955).
12. SCHRIEFFER J. R., *Phys. Rev.* **97,** 641 (1955).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 33–36. Printed in Great Britain.

STUDIES OF AUGER ELECTRONS EJECTED FROM GERMANIUM BY SLOWLY MOVING POSITIVE IONS

H. D. HAGSTRUM

Bell Telephone Laboratories, Murray Hill, New Jersey

Abstract—Measurements have been made of total yield and kinetic energy distribution of Auger electrons ejected from a germanium surface by singly charged noble gas ions. By this means, surface changes have been monitored during procedures designed to clean the surface. Ion bombardment yielded a reproducible surface which showed the Auger characteristics to be expected for the clean surface. Heating at lower temperatures for days and for a total of about 3 hr at 1170°K (40° below the melting point) prior to any sputtering only partially cleaned the surface. From the kinetic energy distribution of the Auger electrons some of the principal characteristics of the state density function in the valence band of germanium have been determined.

VOL. 14 1960

1. INTRODUCTION

THE study of the interaction of slowly moving noble gas ions with solid surfaces is extended in this work to germanium. Previous investigations for metals[1] and for silicon[2] have already been reported. The phenomenon studied is the release of electrons from the solid in the Auger-type neutralization of the ion at the surface. The experimental results are very sensitive to the state of the surface as regards adsorbed impurities and are used to evaluate procedures for cleaning the germanium surface. The results also depend critically upon the distribution of filled electronic states in the solid. The determination of some of the principal features of the state density function in the valence band of germanium will be discussed briefly.

The results concerning initial cleaning of the germanium surface are presented in sufficient detail here. However, the report on the results for the atomically clean surface, its interaction with gas and thermal restoration, and the theoretical interpretation of the results is sketchy and will be expanded in other publications.

2. EXPERIMENT

The electronic process with which we are concerned is the two-electron, Auger-type neutralization of the ion as it comes close to the germanium surface. One electron from the germanium valence band tunnels through the barrier between solid and ion to the ground state of the atom. The second electron picks up the excess energy of this process and becomes a fast internal secondary which may or may not leave the solid. The experiment consists of measuring the total yield and the kinetic energy distribution of those excited electrons which leave the solid.[1] Here we shall be concerned principally with results for He^+ ions on the (111) face of germanium.

The total yield data in Fig. 1 are presented as the quantity γ_i, which is the number of electrons released per incident ion, as a function of ion kinetic energy. The several characteristics shown will be discussed in Section 3. The kinetic energy distributions are obtained from point by point differentiation of retarding potential data. Several such plots are shown in Fig. 2.

The germanium target could be heated by passing current directly through it. The target surface could also be subjected to uniform ion bombardment. An arc was run between an auxiliary heated filament and the Auger electron collector in neon gas at a pressure near 10^{-2} mm Hg (not accurately measured). When the target was operated at a negative potential of 100 V relative to the surrounding arc plasma it was bombarded by an ion current density of 3–5 mA cm^{-2}. Since the average sputtering time was 60 sec, the surface was struck by about 0·24 C cm^{-2} or about

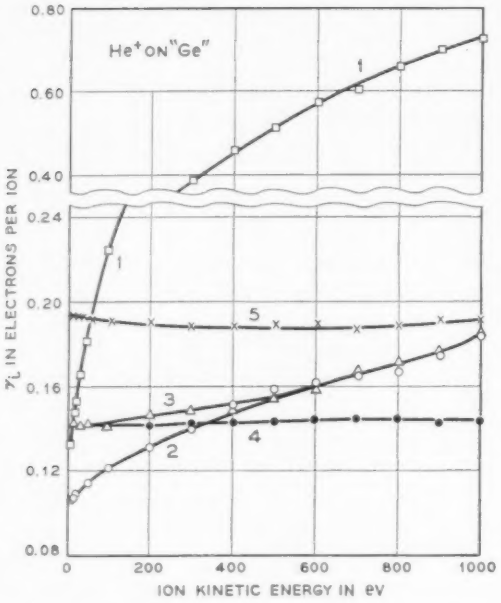


FIG. 1. Total yield, γ_i, of Auger electrons ejected by He^+ ions from the (111) germanium surface in various conditions plotted against incident kinetic energy of the ion.

10^{18} ions cm^{-2} per sputtering procedure. If the material sputtered were all germanium this is sufficient to remove about 100 monolayers.[3]

3. CLEANING OF THE GERMANIUM SURFACE

Measurements of the Auger characteristics γ_i and $N_0(E_k)$ were made for the germanium surface after etching, rinsing, and installation in the experimental tube. These data are shown in Figs. 1 and 2 as curve (1). These characteristics differ radically from those attributable to the atomically clean surface. Measurements were also made on the surface as it was heated in an attempt to clean it. The heating and data-taking schedule is as follows:

Apparatus baked at temperatures between 100° and 210°C for total time of about 200 hr. Target not heated above these temperatures. Data for curves (1) of Figs. 1 and 2 taken during this interval.

Calibration of target temperature versus heating current during which target was heated to temperatures as high as 1150°K for brief intervals (30 sec.).

Target heated to 1125°K for 4 hr.

Apparatus filled with dry N_2 at atmospheric pressure for a minor glass blowing operation, repumped, and baked again at 110°C for about 100 hr. Data taken for curves (2) of Figs. 1 and 2.

Target heated to temperatures below 940°K for 1 hr 48 min. Background pressure now about 2×10^{-9} mm Hg.

Target heated to 940–980°K for 69 hr 25 min.

Target heated to 1000–1090°K for 109 hr 40 min.

Target heated to 1170°K for 10 min. Data taken for curves (3) of Figs 1 and 2.

Target heated to 1170°K for an additional 1 hr 13 min. Data for curve (4) of Fig. 1 taken, no change observed from curve (3) of Fig. 2.

Target heated at lower temperatures and for a total of 1 hr 35 min more at 1170°K producing no change in γ_i or $N_0(E_k)$ characteristics.

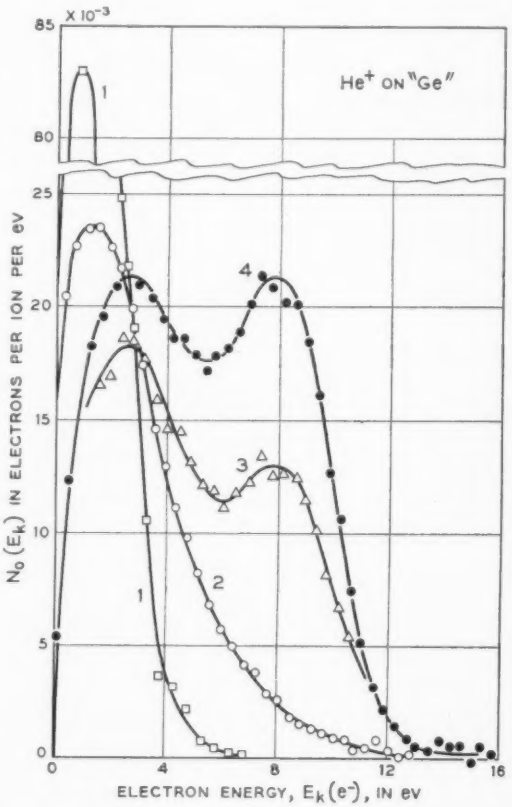


FIG. 2. Kinetic energy distribution, $N_0(E_k)$, of electrons ejected by 10 eV He^+ ions from the (111) face of germanium in various conditions.

One sputtering at about 4 mA/cm² intensity for 90 sec (no heating) produced change from curve (4) to curve (5) of Fig. 1 and from curve (3) to curve (4) of Fig. 2.

Heating to 1170°K for about 3 hr more and four additional sputterings produced no further change in γ_i and $N_0(E_k)$. Background pressure below 2×10^{-9} mm Hg.

The Auger characteristics for the bombarded and annealed surface are those to be expected for the atomically clean surface. The $N_0(E_k)$ distribution is what is expected from the known general form of the state density in the valence band (Section 4). The surface is reactive to oxygen and a drastic change similar to that found for silicon[4] accompanies the adsorption of a monolayer of oxygen. The oxygenated surface gives a $N_0(E_k)$ distribution much like curve (2) of Fig. 2. Furthermore, the Auger characteristics for sputtered germanium are like those for heated silicon for which there is considerable evidence that the surface is atomically clean.[4] It is true that a theoretical fit could also be made to curve (3) of Fig. 2, but this curve cannot be representative of the clean surface since sputtering changes it.

The present work thus corroborates the findings of FARNSWORTH, SCHLIER and others[5] that ion bombardment and annealing will produce an atomically clean germanium surface. It also corroborates the findings of these workers that heating of germanium just below the melting point does not produce the same surface as ion bombardment followed by anneal. ALLEN[6] also finds it impossible to produce the field emission pattern for the atomically clean germanium tip by heating alone.

It should be pointed out that the Auger ejection process is apparently not as sensitive as slow electron diffraction to disorder among the atoms of an atomically clean surface. Thus no essential difference was observed between surfaces which were bombarded only and those annealed after bombardment. However, the Auger characteristics do permit one to estimate the amount of coverage of the surface remaining after heating alone. Comparison of the bombarded curve (4) of Fig. 2 and the heated curve (3) of Fig. 2 with the $N_0(E_k)$ distribution for the oxygenated surface—not shown but like curve (2) of Fig. 2—indicates that the surface after heating has *the equivalent of* a sizable fraction of a monolayer of oxygen remaining on its surface.

LAW[7], using physical adsorption of CO, N_2 and H_2 on germanium, and ALLEN and FOWLER[8], measuring contact potential differences between various crystal faces, found little difference between germanium surfaces which had been heated for long periods to high temperature and those which had been bombarded. It is possible that these experiments are not sensitive to the contamination remaining on the heated surface. There is agreement on the desirability of prolonged heating at lower temperatures to release body gas from the germanium before attempting to clean the surface. There is evidence in the present work, in general agreement with much other work, that a germanium surface, once having been atomically clean, can be restored to this condition after oxygen exposure by heating to temperatures in the range 600–800°K.

4. STATE DENSITY IN THE VALENCE BAND OF GERMANIUM

Theoretical interpretation of the experimental results has been attempted only for the atomically clean surface. For this condition the incoming ion interacts with the valence band electrons directly and not with electrons associated with foreign adsorbed atoms. The theory is similar to that for metals.[9] It has been discussed briefly for semiconductors in Ref. (2) and details will be reported elsewhere. Suffice it to sketch briefly the structure of the theory here. The energy distribution, $N_i(\epsilon_k)$, of excited Auger electrons *inside* the solid is proportional to the integral transform $\int N_v(\epsilon+\Delta)\,N_v(\epsilon-\Delta)\,d\Delta$ of the state density function, $N_v(\epsilon)$, in the valence band. The integral is taken over appropriate limits and the transformation from ϵ to ϵ_k effected. The two valence band electrons which become involved in the process have initial energies in the valence band of $(\epsilon+\Delta)$ and $(\epsilon-\Delta)$. The external energy distribution, $N_0(E_k)$, is obtained from $N_i(\epsilon_k)$ by multiplying by the probability of electron escape over the surface barrier, $P_e(\epsilon_k)$. E_k, the energy outside the solid, is the energy inside, ϵ_k, minus the total barrier height, ϵ_0. Account must be taken of the s versus p character of the valence band wave functions.[2] Broadening of the energy distribution by both the variation of the effective ionization energy of the atom near

the solid and the Heisenberg uncertainty principle must also be included.[9] The fit of the theory to the experimental results involves the determination of several parameters whose effects are very nearly orthogonal to one another. Among these are parameters which specify some of the features of the density of states in the valence band.

The form of the experimental kinetic energy distribution of Auger electrons released from germanium through the atomically clean surface is thus determined in large measure by the state density function in the valence band. The four valence electrons are known to be distributed as follows. Two electrons of p character lie in degenerate bands which are relatively narrower than the band over which the two remaining electrons are spread. The wave functions of these latter two electrons vary from pure s character at the bottom to pure p character at the top. Neglecting spin-orbit interaction (unobservable in the present work) all bands have the same upper energy limit.

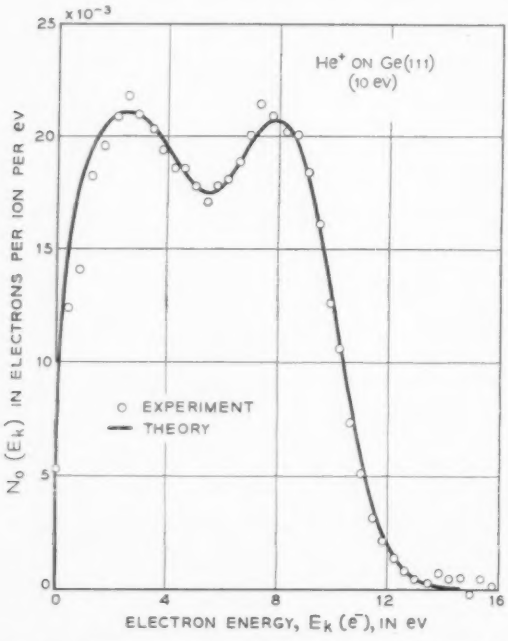


FIG. 3. Comparison of experimental and theoretical $N_0(E_k)$ distributions of electrons ejected by 10 eV He^+ ions from Ge (111). The experimental points are those of curve (4) of Fig. 2. The curve is a theoretical one obtained from the numerical calculation carried out on the IBM 704 computer.

By fitting theory to experiment it is possible to get a quite accurate value for the width of the degenerate p bands. A less accurate estimate of the total width of the valence band is also possible.

The theory[2, 9] of the Auger neutralization process has been programmed for the IBM 704 computer and a fit of the experimental data achieved by variation of the several parameters in the problem. The result for 10 eV He^+ ions on the (111) face of clean germanium is shown in Fig. 3. Considering the possible limitations of the theoretical model the fit is considered to be very good. A considerably more extensive publication than the present one will be required to discuss all the elements of this problem.

The theoretical fit yields 4·4 eV as the width of the degenerate p bands. This is only about 10 per cent narrower than that obtained for silicon.[2] The fit was accomplished for a total valence band width of 16 eV. However, at present writing, variation of this parameter has not been completely explored. Any change in the total width will have no effect on the 4·4 eV width of the degenerate p bands, however.

Acknowledgement—It is a pleasure to acknowledge the help of Mr. C. D'AMICO in the operation of the experimental apparatus; of Mrs. WANDA MAMMEL, who programmed the theory for the IBM 704 computer; and of Miss M. C. GRAY, who provided helpful consultation on mathematical problems.

REFERENCES

1. HAGSTRUM H. D., *Phys. Rev.* **96,** 325 (1954), tungsten; *Phys. Rev.* **104,** 672 (1956), molybdenum.
2. HAGSTRUM H. D., *J. Phys. Chem. Solids* **8,** 211 (1959).
3. LAEGREID N., WEHNER G. and MECKEL B., *J. Appl. Phys.* **30,** 374 (1959).
4. ALLEN F. G., EISINGER J., HAGSTRUM H. D. and LAW J. T., *J. Appl. Phys.* **30,** 1563 (1959).
5. FARNSWORTH H. E., SCHLIER R. E. and BURGER R. M., *J. Appl. Phys.* **26,** 252 (1955).
   SCHLIER R. E. and FARNSWORTH H. E., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 3. University of Pennsylvania Press, Philadelphia (1957).
   FARNSWORTH H. E., SCHLIER R. E., GEORGE T. H. and BURGER R. M., *J. Appl. Phys.* **29,** 1150 (1958).
6. ALLEN F. G., private communication.
7. LAW J. T., *J. Phys. Chem.* **59,** 543 (1955).
8. ALLEN F. G. and FOWLER A. B., *J. Phys. Chem. Solids* **3,** 107 (1957).
9. HAGSTRUM H. D., *Phys. Rev.* **96,** 336 (1954).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 37–42. Printed in Great Britain.

IMPURITY CONDUCTION OF CLEANED GERMANIUM SURFACES AT LOW TEMPERATURES

A. KOBAYASHI, Z. ODA, S. KAWAJI,* H. ARATA and K. SUGIYAMA

Electrical Communication Laboratory, Tokyo, Japan

Abstract—The conductivity and Hall effect of cleaned germanium surfaces have been measured in the temperature range between 1·7°K and 300°K. The surface was cleaned by Joule heating at ~800–850°C in ultra-high vacuum for ~20–30 hr and was then annealed at 650°C for several hours.

Results obtained with ~20–40 Ω-cm *p*-type samples show that the Hall coefficient has a steep maximum at ~10–20°K and the resistivity becomes almost flat below this temperature range. This character is similar to that of bulk impurity conduction found by HUNG *et al.*, but disappears on exposure to air which causes marked increases in both the Hall coefficient and resistivity at low temperatures. This impurity conduction is attributed to the surface imperfections introduced by the cleaning treatment of the surface. The mobility and concentration of carriers in this surface conduction are estimated.

VOL. 14 1960

1. INTRODUCTION

RECENT investigations[1, 2] on cleaned germanium surfaces indicate that there are a large number of states localized on the surface, predominantly related to the unpaired bonds on the germanium surface. The density of these surface states is estimated to be as high as 10^{13} states/cm^2 which behave essentially like acceptor levels making the surface strongly *p*-type even on an *n*-type crystal.

In connection with these properties of the cleaned germanium surface, it may be expected that the surface states should give rise to a kind of impurity conduction at very low temperatures. In this paper we will report the work undertaken along this line.

2. EXPERIMENTAL METHODS

The samples used are slices of an *n*-type 40 Ω-cm germanium single crystal with four side probes as shown in Fig. 1(a). The surfaces are parallel to a (111) face. Polishing, etching in CP-4 and subsequent rinsing were carried out following the ordinary procedures. In order to avoid plastic deformation during heat-treatment, very thin flexible tantalum ribbons (0·03 cm in thickness and 2 mm in width) were welded to the crystal at both ends and at the four probes. This assembly was mounted on molybdenum pinch leads and sealed in a glass bulb as shown in Fig. 1(b).

The bulb was evacuated to ultra-high vacuum by a Hickman oil diffusion pump or a mercury diffusion pump through two liquid oxygen traps.

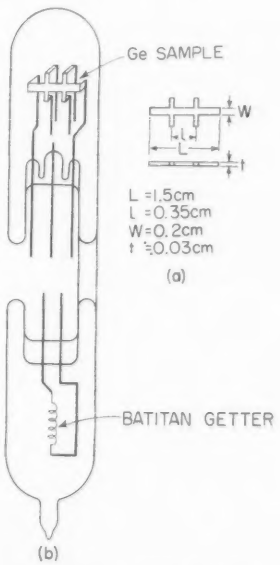


FIG. 1. Schematic diagram of (a) Ge slice and (b) the experimental tube.

* Gakushuin University, Tokyo, Japan.

After preliminary bake-out at 430°C for 3 hr, the Batitan getter* was outgassed below 1000°C. In this process barium carbonate was decomposed into the oxide and CO_2 which was completely evacuated. Then the system was again baked out thoroughly. When the pressure was reduced down to 10^{-9} mm Hg, the sample was heated by passing a.c. current through it. The maximum temperature at outgassing was 850°C which was measured with an optical pyrometer using the emissivity data of ALLEN[3].

At the initial stage of outgassing the pressure was as high as 2×10^{-7} mm Hg, but decreased to 10^{-9} mm Hg over the total outgassing period of 20–30 hr. Before cooling, the sample was always annealed at about 650°C for at least 30 min.

After flashing the Batitan getter, the bulb was sealed off from the pumping system. Outgassing and annealing of the sample was repeated again in order to remove surface contamination possibly caused by gases liberated from glass at the time of sealing-off.

The resistivity, ρ, Hall coefficient, R, and magnetoresistance, ρ_H, of the cleaned surfaces were measured between 1·7° and 300°K. Subsequently, the sample was exposed to atmospheric air for several days or to wet air for a short time and then re-evacuated to 10^{-6} mm Hg. On this air-exposed surface the same measurements were repeated.

3. RESULTS

The above mentioned heat-treatment in a high vacuum converted the material from n-type of 40 Ω-cm into p-type of ~20–40 Ω-cm. Therefore, it was impossible to measure on n-type samples.

Figures 2 and 3 show plots of log ρ vs. $1/T$ and log R vs. $1/T$ for the samples with cleaned and air-exposed surfaces. Fig. 4 shows the temperature dependency of Hall mobility.

As seen easily from Figs. 2 and 3, the cleaned surface shows a steep maximum in the Hall curve between 10 and 20°K and the slope of the resistivity curve becomes almost flat below this temperature range. However, these low-temperature anomalies disappear on exposure to air which causes marked increases in R and ρ.

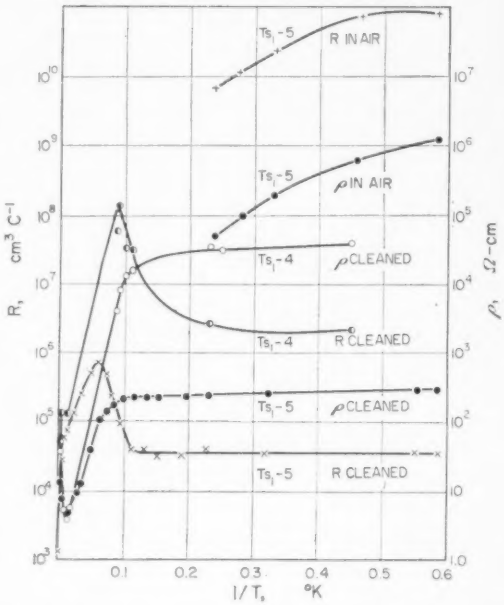


FIG. 2. Plots of log ρ vs. $1/T$ and log R vs. $1/T$ for the cleaned surfaces of Ts_1-4 and -5.

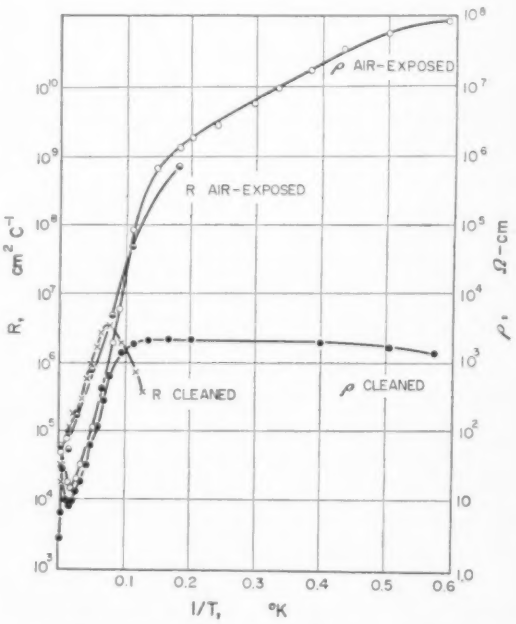


FIG. 3. Plots of log ρ vs. $1/T$ and log R vs. $1/T$ for the cleaned and air-exposed surfaces of Ts_1-6.

* A reaction type barium getter using a titanium coil as a reducing agent for barium carbonate.

Table 1. *Summary of the results obtained for germanium with cleaned surfaces*

	Ts_1-3	Ts_1-4	Ts_1-5	Ts_1-6
(300°K) Ω-cm	6·6	31	13	38
R(300°K) cm³/C	$1{\cdot}6\times10^{3}$	$1{\cdot}2\times10^{5}$	3×10^{4}	7×10^{4}
N_A-N_D cm^{-3}		$5{\cdot}2\times10^{13}$	$2{\cdot}1\times10^{14}$	9×10^{13}
E_A eV		0·0086	0·005	0·0062
N_A cm^{-3}				$1{\cdot}37\times10^{16}$
N_D cm^{-3}				$1{\cdot}36\times10^{16}$
N_{scatt} cm^{-3}		1×10^{15}	$3{\cdot}8\times10^{15}$	$1{\cdot}4\times10^{16}$
R(2°K) cm³/C	$1{\cdot}4\times10^{4}$	$2{\cdot}7\times10^{6}$	$3{\cdot}4\times10^{4}$	$<2{\cdot}8\times10^{5}$*
p(2°K) cm^{-3}	$4{\cdot}5\times10^{14}$	$2{\cdot}3\times10^{12}$	$1{\cdot}8\times10^{14}$	$>2{\cdot}2\times10^{13}$*
ρ(2°K) Ω-cm	52	4×10^{4}	$2{\cdot}3\times10^{2}$	$2{\cdot}2\times10^{3}$
t (thickness) cm	0·033	0·03	0·03	0·03
ρ_s(2°K) Ω	$3{\cdot}15\times10^{3}$	$2{\cdot}7\times10^{6}$	$1{\cdot}5\times10^{4}$	$1{\cdot}5\times10^{5}$
p_s cm^{-2}	$7{\cdot}4\times10^{12}$	$3{\cdot}5\times10^{10}$	$2{\cdot}7\times10^{12}$	$>3{\cdot}3\times10^{11}$*
μ_s(2°K) cm²/V. sec	270	68	148	<127*

Keys to symbols: $p_s = p(2°K)\times(t/2)$, $\rho_s = \rho(2°K)\times(2/t)$, $\mu_s = R(2°K)/\rho(2°K)$.

* These values are obtained at 8·3°K, where R is decreasing with temperature.

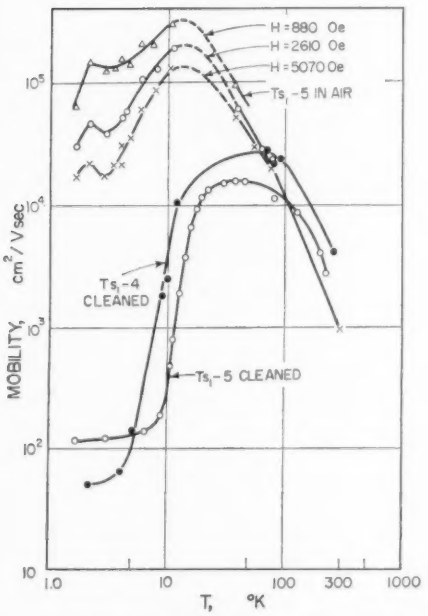


FIG. 4. Temperature dependency of mobility for Ts_1-4 and -5.

The Hall mobility at these low-temperature anomalies is very low as shown in Fig. 4, but increases by three orders of magnitude after air-exposure in good agreement with the extrapolated values from the high-temperature range of cleaned surface.

Table 1 lists the summary of the data obtained with cleaned surfaces. In this table, N_{scatt} refers to the concentration of ionized scattering centers calculated from the mobility at 90°K according to HERRING and BROOK's formula for ionized impurity

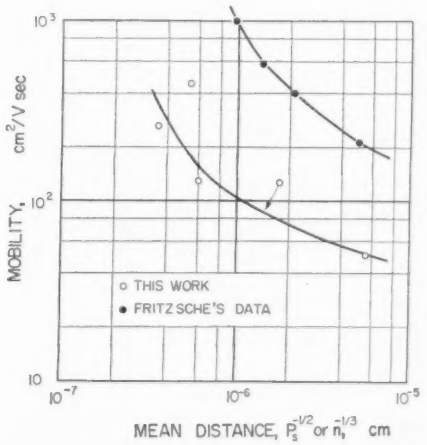


FIG. 5. Surface mobility as a function of the mean distance between neighbouring surface carriers. For comparison, the mobility of bulk impurity conduction derived from FRITZSCHE's data is also plotted. The mean distance is given as square root of p_s and cube root of n, respectively.

scattering. The concentrations of donors, N_D, and acceptors, N_A, and the activation energy for conduction, E_A, have been determined by DEBYE and CONWELL's method.

$\rho_s(2°K)$ refers to the surface resistivity per cm^2 at 2°K and is defined as $\rho_s = \rho(2°K) \times (2/t)$, where t is the thickness of the sample. p_s is the concentration of surface carriers which result in the anomalous conduction and will be given by

$$p_s = p(2°K)\, t/2,$$

where $p(2°K)$ is the carrier concentration directly obtained from the Hall coefficient. $\mu_s(2°K)$ refers to the surface mobility having the same dimension as the bulk one and so is given by $R(2°K)/\rho(2°K)$ As illustrated in Fig. 5, μ_s is increased with the mean distance between neighbouring surface carriers.

4. DISCUSSION

The low-temperature anomaly of conduction observed at the cleaned surfaces disappears as soon as they are exposed to air. This change, however, does not proceed any further by prolonged oxidation at room temperature. This fact suggests that the anomalous conduction takes place on the surface only or in such a thin surface layer as to be oxidized immediately after air-exposure.

Therefore, we assume two different mechanisms of conduction, the one taking place in the bulk and the other on the surface. Above the Hall maximum temperature the bulk conduction dominates, whereas below that temperature the surface conduction exhibits itself instead of the bulk one.

From Table 1, the concentration of surface carriers, p_s, is found to lie between 10^{10} and 10^{13}/cm^2. Taking this fact into account three models shown in Fig. 6 may be tentatively constructed for the cleaned surface:

Model 1.* When the surface is negatively charged due to unknown causes, a space charge layer will be built up by positive holes near the surface giving rise to the degeneration of carriers. Consider the idealized potential well in which the potential is a linear function of the distance from the surface as shown in Fig. 7. Since the surface electric field is very strong, the kinetic motion of

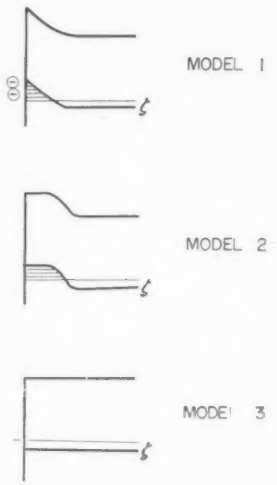


FIG. 6. Three models showing the energy bands at the cleaned surface of p-type germanium, when (1) surface is negatively charged, (2) imperfections are segregated to the surface layer, and (3) surface states are situated close to the top of the valence band.

a hole normal to the surface is quantized and its energy is given by the following discrete levels,[5]

$$E_n = \frac{1}{2}\left\{\frac{(3\pi e\alpha\hbar)^2}{m}\right\}^{1/3}(n+\tfrac{3}{4})^{2/3},\ n = 0,1,2\ldots$$

where $\alpha = 4\pi e p_s/\epsilon$ is the electric field strength normal to the surface, and ϵ the dielectric constant of the crystal. Energy levels for this model were calculated as a function of p_s. The results are given in Table 2 in practical units.

From Table 2 it can be seen that the mean free path z_0 is of the order of 10^{-7}cm and decreases as p_s increases. This model can possibly explain the

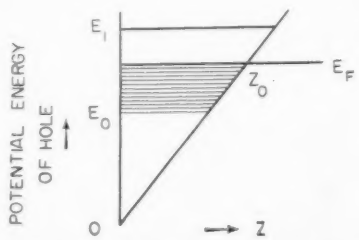


FIG. 7. Idealized potential well showing discrete levels, E_0, E_1 and Fermi level E_F.

* This model is somewhat different from that of FREDERIKSE *et al.*[4] proposed in the case of Mg_2Sn.

Table 2. Variation of parameters of space charge layer when the surface carrier density is changed

p_s (cm^{-2})	α (V/cm)	E_0 (eV)	$E_F - E_0$ (eV)	Z_0 (cm)
1×10^{10}	$1{\cdot}13\times10^{3}$	$1{\cdot}17\times10^{-3}$	$3{\cdot}16\times10^{-5}$	$1{\cdot}6\times10^{-6}$
1×10^{11}	$1{\cdot}13\times10^{4}$	$8{\cdot}25\times10^{-3}$	$3{\cdot}16\times10^{-4}$	$7{\cdot}3\times10^{-7}$
1×10^{12}	$1{\cdot}13\times10^{5}$	$3{\cdot}84\times10^{-2}$	$3{\cdot}16\times10^{-3}$	$3{\cdot}4\times10^{-7}$
1×10^{13}	$1{\cdot}13\times10^{6}$	$1{\cdot}77\times10^{-1}$	$3{\cdot}16\times10^{-2}$	$1{\cdot}6\times10^{-7}$

small magnitude of μ_s in the anomalous conduction. but contradicts the data shown in Fig. 5. Thus, it is impossible to accept this model.

Model 2. This model shows the impurity conduction through a surface layer with a large number of imperfections possibly segregated by the heat-treatment of the sample. Since an anomalous conduction has been found along a grain boundary of germanium crystal,[6] we should pay attention to this model.

On the cleaned surface we have observed thermal etch pits of the order of $10^6/cm^2$ in density. However, these pits are not exclusively attributable to the dislocations because they have a flat bottom and a chemical etching can easily erase them, leaving much fewer chemical etch pits of $10^4/cm^2$ in density.

Assume the thickness of the surface layer which would suffer oxidation on exposure to air to be 100 Å. If the carriers arise from dangling bonds situated at every 4 Å in distance along a dislocation line, it will be required for $p_s = 10^{12}/cm^2$ that there should be so high a density of dislocation lines as $4\times10^{10}/cm^2$.* However, this value of density cannot compare with that of the samples used.

However, in spite of much care taken during the mounting and heat-treatment of the sample, there remains another possibility that some impurities would diffuse into the crystal and segregate at the surface giving rise to a highly doped surface layer. Tantalum would not dissolve in the crystal according to the experiment of MAYBURG[7]. It is conceivable that metals with high binding energy are hard to segregate at the surface.[8]

* The density of dislocation lines can be estimated by p_s/(thickness of surface layer)/(number of dangling bonds per dislocation line) $= 10^{12}/10^{-6}/(4\times10^{-8}) = 4\times10^{10}/cm^2$.

Model 3. We consider the case where unpaired bonds attributed to the Tamm–Shockley levels and surface steps on the cleaned surface could cause the impurity conduction. Let $2r_s$ be the average distance between neighbouring carriers on the surface, the variation of r_s with p_s is given as follows:

$p_s(cm^{-2})$	r_s(cm)
10^{10}	5×10^{-6}
10^{11}	$1{\cdot}6\times10^{-6}$
10^{12}	5×10^{-7}

Comparing these values of surface carrier concentration, p_s, with those of the bulk impurity conduction by FRITZSCHE[9], one can easily find that $p_s = 10^{11}/cm^2$ corresponds to the "intermediate concentration range" and $p_s = 10^{12}/cm^2$ to the "high concentration range". Therefore, the overlapping of wave functions of neighbouring surface carriers is thought to be sufficient to give rise to the anomalous conduction. Furthermore, the wave function of carriers should extend considerably as their mobility increases, thus the level may be situated very close to the top of valence band. This model is not inconsistent with the fact that there is a strongly p-type inversion layer on the cleaned surface of n-type crystals.

5. CONCLUSION

The low-temperature anomalies in the Hall coefficient and the resistivity of p-type germanium with a cleaned surface seems to be ascribable to the surface levels close to the top of valence band or an extremely thin surface layer which would be highly doped with unknown imperfections.

Acknowledgements—The authors would like to acknowledge the supply of liquid helium and hydrogen by the Research Institute for Iron, Steel and Other Metals and by the Electrotechnical Laboratory. They are indebted to T. TAKAHASHI and his colleagues for the preparation of the crystals and to Y. UEMURA, J. OKADA and R. SUGITA for many stimulating discussions.

REFERENCES

1. HANDLER P., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 23. University of Pennsylvania Press, Philadelphia (1957).
   MISSMAN R. and HANDLER P., *J. Phys. Chem. Solids* **8,** 109 (1959).
2. KOBAYASHI A. and KAWAJI S., *J. Phys. Soc. Japan.* **12,** 1057 (1957).
3. ALLEN F. G., *J. Appl. Phys.* **28,** 1510 (1957).
4. FREDERIKSE H. P. R. *et al.*, *Phys. Rev.* **103,** 67 (1956).
5. FLÜGGE S. und MARSHALL H., *Rechenmethoden der Quanthentheorie* Bd. I, S.70. Springer, Berlin (1952).
6. REED B., WEINREICH D. A. and MATARÉ H. F., *Phys. Rev.* **113,** 454 (1959).
7. MAYBURG S., *Phys. Rev.* **95,** 38 (1954).
8. HANNAY N. B., *Semiconductors* p. 156. Reinhold, New York (1959).
9. FRITZSCHE H., *J. Phys. Chem. Solids* **6,** 69 (1958).

VOL
14
196

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 43–50. Printed in Great Britain.

SECTION II. NEW TECHNIQUES

APPLICATION OF MAGNETIC RESONANCE TECHNIQUES TO THE STUDY OF SEMICONDUCTOR SURFACES

G. K. WALTERS

Central Research Laboratories, Texas Instruments Incorporated, Dallas, Texas

VOL. 14 1960

Abstract—After a brief discussion of the magnetic resonance phenomenon, experimental methods and sensitivity considerations, there follow sections on the application of electron and nuclear paramagnetic resonance techniques respectively to the study of solid surfaces. Attention is given to the pioneering work done in these fields. The various types of surface information which one might hope to obtain by resonance investigations are outlined, and some preliminary new experimental results are presented.

1. INTRODUCTION

RECENT experimental investigations conducted in a number of laboratories have indicated that nuclear and electron paramagnetic resonance techniques can be applied successfully to the study of solid surfaces.[1–6] As applied in particular to semiconductor surfaces, these techniques offer potentially important new methods in a field of study where experimental approaches are at present severely limited.

In this paper, we will review some of the pioneering work which is illustrative of the types of surface information one may obtain by use of the resonance techniques. We will also present recent results obtained in our own laboratory, by application of both nuclear and electron paramagnetic resonance methods to semiconductor surfaces.

It is of importance to consider briefly some of the characteristics of the magnetic resonance methods which make them potentially valuable to the surface physicist. As HANDLER points out,[7] the more frequently used conventional experimental methods for semiconductor surface study suffer from the common disadvantage that measurements are made in regions adjacent to the surface, rather than in the surface layer itself. Since the resonance techniques in some cases may be applied directly to the study of the surface film, it would appear that the conventional techniques and the resonance techniques may complement one another quite nicely. In fact, application of both conventional and resonance methods to the same samples may be especially fruitful.

From the resonance signal intensity and shape function, one can determine in most cases the density of each species of paramagnetic center present. As will be discussed in Sections 3 and 4, information often can be obtained regarding the degree of localization of the paramagnetic centers in question. Especially in the case of nuclear magnetic resonance, one has the additional desirable characteristic that there is never any question as to which nuclear species is being investigated; each paramagnetic nucleus has its own unique resonance condition.

In the following section, a brief outline will be given of the conditions required for the observation of magnetic resonance, and the pertinent relaxation parameters will be discussed briefly. In Sections 3 and 4 the application of electron and nuclear magnetic resonance techniques to surface studies will be discussed.

2. THE MAGNETIC RESONANCE PHENOMENON

There are available a number of good books[8, 9] and comprehensive review articles[10, 11] which discuss in detail the magnetic resonance theory,

experimental methods, and sensitivity considerations.[12] We will review here only briefly the salient features of the magnetic resonance phenomenon as it may be applied to the study of solid surfaces.

(a) *Theory*

For the electron paramagnetic resonance phenomenon to occur, the material under investigation must contain a number of paramagnetic atoms, ions, free radicals, or other centers, each of which possesses a resultant angular momentum and permanent magnetic moment. Similarly, in the case of nuclear magnetic resonance, nuclei with non-zero angular momentum and permanent magnetic moments must be present.

If a sample containing paramagnetic centers of resultant angular momentum J is placed in a constant external magnetic field H_0, the energy levels corresponding to the different allowed spatial orientations of J are given by $Mg\beta H_0$. M is the magnetic quantum number and β is either the Bohr magneton or the nuclear magneton, depending on whether we speak of electron or nuclear paramagnetic resonance.

g is the spectroscopic splitting factor. For free electrons its value is $g_e = 2{\cdot}0023$; deviations of resonance line positions from $g = g_e$ often occur and are accounted for by the presence of orbital angular momentum and/or interactions of the paramagnetic center with its environment. For nuclear paramagnetism, g is a number characteristic of the nuclear species.

If the sample is excited by electromagnetic radiation of frequency ν applied at right angles to H_0, the paramagnetic centers will absorb energy from the radiation field when the resonance condition

$$h\nu = g\beta H_0 \qquad (1)$$

is satisfied, where the selection rules for allowed transitions are $\Delta M = \pm 1$. For normal laboratory magnetic fields, the frequency ν is typically in the radio-frequency spectral region for nuclear magnetic resonance, and in the microwave region for electron paramagnetic resonance.

For a system in thermal equilibrium, the lower the energy of the state the higher will be its population. Since the probabilities for stimulated emission and absorption between two states are equal, at resonance there will be a net transfer of energy from the radiation field to the spin system of the sample, corresponding to a shift toward more equal populations of the energy levels involved. The rate P at which energy is absorbed from the radiation field is

$$P = \tfrac{1}{2}\omega H_1^2 \chi'' \qquad (2)$$

where H_1 is the amplitude of the magnetic field alternating with angular frequency $\omega = 2\pi\gamma$ and χ'' is the imaginary part of the complex susceptibility.

The resonance phenomenon may be further characterized by the spin–lattice relaxation time T_1 and the spin–spin relaxation time T_2.

As discussed above, at resonance the radiation field tends to upset the thermal equilibrium distribution of spins among the various energy levels. The new equilibrium, in the presence of resonance radiation, will depend on the spin–lattice relaxation time T_1. T_1 characterizes the return to thermal equilibrium of a spin system, starting from an initial non-equilibrium state.

The spin–spin relaxation time T_2, in many instances, may be considered as a measure of the inverse resonance line width. This description is adequate for simple isolated spin systems in which the resonance line width arises primarily either from interactions between the various resonant spins, or from a broadening of the magnetic energy levels due to the finite lifetime of a spin state. In the latter case, called T_1 broadening, $T_2 \simeq T_1$.

In more complex spin systems, "inhomogenous" broadening[13] of the resonance line may result from a spread in g-values arising from local perturbations of the paramagnetic center. In this case, we may think of the resonance line as being composed of a large number of narrower "homogeneously" broadened resonance lines or spin packets, which are unresolved. In this case, the T_2 for a spin packet is the relaxation time of interest.

(b) *Experimental techniques*

In Fig. 1 we show schematically a simple nuclear magnetic resonance apparatus which illustrates the technique principles. The sample is contained within the coil of a tuned circuit which is placed in a magnetic field H_0.

VOL. 14 1960

Radio-frequency energy is fed to the circuit from a constant current source. Under these conditions, the r.f. voltage developed across the coil is directly proportional to the quality factor Q of the tuned circuit. At magnetic resonance, the

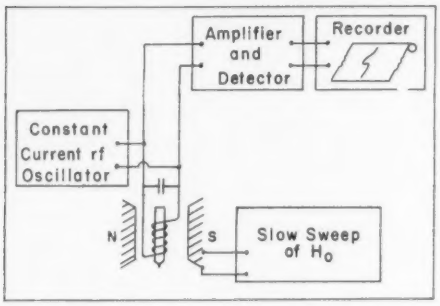


FIG. 1. Simplified schematic diagram of nuclear magnetic resonance spectrometer.

absorption of power by the sample from the r.f. field has the effect of lowering the circuit Q, hence of lowering the voltage across the coil.

In practice, the radio frequency is fixed and the magnetic field is swept slowly through the resonance value. Hence, the power absorbed by the sample is recorded at fixed frequency as a function of applied external magnetic field.

For electron paramagnetic resonance, due to the much higher frequencies employed, one generally goes over to microwave techniques and replaces the r.f. tank circuit of Fig. 1 with a resonant cavity. However, spectra are recorded in the same manner as described above for nuclear magnetic resonance.

As will be discussed in more detail in Section 4, in the study of adsorbed species by nuclear resonance techniques it is the relaxation parameters T_1 and T_2 which are of primary interest. The c.w. techniques described above are in this case not nearly so convenient as are the spin-echo pulse techniques first used by HAHN[14, 15]. Without discussing the technique further, we will describe the form in which the raw data is taken, and how it can be reduced to yield values for T_1 and T_2.

In measuring T_2, we excite the sample with pulses of r.f. energy at times $t = 0$ and $t = \tau$, as shown in Fig. 2a. At time 2τ, the spin echo signal arrives (Fig. 2b) and is picked up by means of the same coil through which the r.f. pulses were supplied. The echo amplitude is proportional to $\exp(-2\tau/T_2)$. Hence by plotting the logarithm of the echo amplitudes as a function of twice the variable τ (the time between pulses), one obtains T_2 from the slope of the resulting straight line.

As shown in Fig. 2c, three pulses are required for a measurement of T_1. The separation τ of the

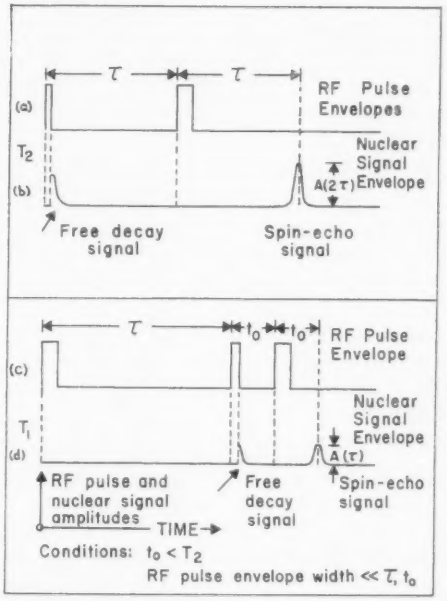


FIG. 2. Schematic diagram of r.f. pulse sequences and resultant spin-echos for determination of relaxation times T_1 and T_2: (a) applied r.f. pulse sequence for T_2 determinations; (b) spin-echo signal proportional to $\exp(-2\tau/T_2)$; (c) applied r.f. pulse sequence for T_1 determination; (d) spin-echo signal proportional to $[1-2\exp(-\tau/T_1)]$.

first two pulses is varied until a null value is obtained for which there is no echo. We then have $T_1 = \tau_{\text{null}}/\ln 2$.

(c) *Sensitivity considerations*

In the study of paramagnetic centers associated with a solid surface, sensitivity considerations become extremely important because of the small number of centers present in the experimental samples.

The power absorbed per unit volume by the sample at magnetic resonance is given by equation

(2). For a sample containing N paramagnetic centers, each with spin I,[10]

$$\chi'' = \frac{\pi\nu g^2\beta^2 N\,[I(I+1)-M(M-1)]}{8kT\,(2I+1)} f(\nu) \qquad (3)$$

for the $M \leftrightarrow (M-1)$ transition. T represents absolute temperature. The function $f(\nu)$ describes the resonance line shape, and is normalized so that

$$\int_0^\infty f(\nu)\,d\nu = 1.$$

One may derive an expression for the minimum detectable χ'' by equating the resonance signal voltage resulting from equation (2) and the noise voltage due to the equivalent resistive impedance of the resonance and detection circuits. Using a receiver with noise figure F and bandwidth $\Delta\nu$,[10]

$$\chi''_{\min} \cong \frac{V}{\pi Q}\left(\frac{kTF\Delta\nu}{P_0}\right)^{1/2} \qquad (4)$$

where V is the effective volume of the resonant cavity or r.f. coil, Q is its quality factor, k is the Boltzmann constant, and P_0 is the power available from the electromagnetic energy source.

In a typical X-band ($\simeq 3$ cm wavelength) electron paramagnetic resonance spectrometer, $V \simeq 3\ \text{cm}^3$, $\Delta\nu \simeq 0{\cdot}1$ c/s, $P_0 \simeq 10^{-3}$ W, $Q \simeq 5000$, and one might expect $F \simeq 10$. Under these conditions, at room temperature $\chi''_{\min} \simeq 10^{-13}/\text{cm}^3$. For a resonance line width (in frequency units) to frequency ratio of 0·01 (typical for electron spin resonance lines from semiconductor surfaces), the value obtained for $\chi''_{\min}$ corresponds to about 10^{12} paramagnetic centers. As is seen from equations (3) and (4), even greater sensitivity can be achieved by lowering the temperature.

From the above considerations, it is apparent that even for a fairly low surface density of electron paramagnetic centers, sensitivity will often not be a problem.

In contrast, the observation of nuclear magnetic resonance from centers associated with the solid surface is always complicated by sensitivity problems. Here one can detect nearly as small a χ'' as is detectable in electron paramagnetic resonance; however, nuclear magnetic moments are of the order of 1000 times smaller than electronic moments and the frequencies used about 300 times lower. Hence, equation (3) tells us that we will need perhaps 10^8 times as many nuclear magnetic moments for detection as we needed electronic moments. Even for finely divided powders, with large surface areas, one needs a relatively high surface density of nuclear magnetic moments if a resonance signal is to be detected.

3. SURFACE STUDIES UTILIZING ELECTRON PARAMAGNETIC RESONANCE TECHNIQUES

In the course of their investigations at liquid helium temperatures on shallow donor states in silicon, FLETCHER *et al.*[16] and later FEHER[6] observed an electron paramagnetic resonance line which is associated with the silicon surface. The line appears when the sample surface is mechanically damaged by sand blasting, lapping, or crushing into small particles. The resonance characteristics are unaffected by chemical treatments which act only to remove, produce, or otherwise alter the surface film present on the silicon. The resonance line disappears, however, upon removal by etching of $\sim 10^{-4}$ cm of the silicon surface, and reappears if the surface is again damaged mechanically.

For silicon samples sandblasted with 600 mesh SiC, FEHER[6] reports a surface density of paramagnetic centers $\sim 3 \times 10^{14}$ per cm^{-2}, a g-value of 2·006, $T_2 \simeq 10^{-8}$ sec (corresponding to a homogeneously broadened line of half-width $\Delta H \simeq 6$G between points of maximum and minimum slope), and $T_1 \simeq 10^{-5}$ sec. The same resonance line is observed in both n- and p-type silicon.

We have performed a few exploratory experiments on the surface resonance characteristics of both silicon and germanium samples, which were mechanically damaged in air by crushing into small particles. An average particle size of about 0·7μ was obtained by violent agitation for 30 min at 1200 c.p.m. in a Spex Industries Model 8000 Mixer-Mill, using a brass container and "Arlcite" ceramic grinding balls (manufactured by Patterson Foundry and Machine Company). Small amounts of brass and ceramic were present in the resulting powders. The ceramic exhibits a weak, asymetric, resonance line at about $g = 2{\cdot}00$, and its small contribution to the semiconductor surface resonance spectrum was taken into account.

For silicon powdered by the above described

process, an intense electron paramagnetic resonance line is observed which has the same characteristics that FEHER observed for lapped or sand-blasted silicon platelets. This resonance line is easily observable at room temperature, and its associated paramagnetic susceptibility exhibits a linear inverse temperature dependence down to 1·3°K, the lowest temperature at which we made measurements.

Germanium powdered by the same process yields a single resonance line at $g = 2{\cdot}024 \pm 0{\cdot}001$, with line width between points of maximum and minimum slope $\Delta H \simeq 50$ G. Between 1·3°K and 4·2°K, the susceptibility is proportional to inverse temperature. The volume density of paramagnetic centers in powdered germanium is about an order of magnitude lower than in powdered silicon of approximately the same particle size.

VOL. 14 1960

On the basis of existing experimental data, it is impossible to draw definite conclusions as to the detailed electronic structure of these paramagnetic centers. The positive sign of $\Delta g = g - g_e$ for both silicon and germanium implies that the centers are holes if they are non-localized, or that, if localized, they are due to electronic shells which are more than half completed.[17] The centers are located either within the semiconductor bulk very near the surface, or in the interface region between semiconductor and surface film.

In silicon, our unsuccessful search for hyperfine structure from the 4·7 per cent abundant Si^{29} isotope indicates that the surface paramagnetic centers are not highly localized on silicon atoms. This view is further supported by the fact that the resonance line is homogeneously broadened.

CLARKE and HOPKINS[18], and BUCK and MCKIM[19] have shown that mechanically damaged germanium surfaces exhibit strong p-type conductivity. Assuming that damaged silicon surfaces show similar behavior, it appears possible that the surface paramagnetic resonance line could arise from the same holes as are responsible for this surface conductivity.

We are unable to say whether the presence of an ambient atmosphere (oxygen in particular) is required during or after the mechanical damaging process in order for the paramagnetic centers to appear. Investigation of samples damaged in vacuum should answer this question and such studies are presently underway in our laboratory.

4. SURFACE STUDIES UTILIZING NUCLEAR MAGNETIC RESONANCE TECHNIQUES

The nuclear relaxation processes described by T_1 and T_2 arise from interactions between the nuclear magnetic moments and the local magnetic field. The local field at the position of the nucleus under study arises from the presence of neighboring magnetic nuclei, and the local field spectrum consists of both static and oscillating components.

The high frequency components of the local field determine the spin–lattice relaxation time T_1. In particular, the Fournier component at the nuclear magnetic resonance frequency induces transitions between the energy levels in much the same way as does an applied electromagnetic field.

The static and low frequency components of the local field add or subtract from the externally applied field so that different nuclei may see different total fields. This results in a broadening of the resonance line, hence in a decreasing T_2.

We may consider two extremes. On the one hand, the local field in a liquid is caused by the rapid Brownian motion of the magnetic nuclei. In this case, the small static and low frequency components of the local field spectrum are ineffective and the resonance line is very narrow. Generally, T_2 is limited by the lifetime of a spin state, i.e.

$$T_1 \cong T_2.$$

On the other extreme, in most solids, the nuclear magnets are rigidly held to fixed lattice sites. In this case, the static component of the local field is large, and the resonance line is broad corresponding to a very small T_2. The high-frequency components of the local field are small and T_1 becomes quite large.

As an illustration, for the hydrogen nuclei in water at room temperature, $T_1 \simeq T_2 \simeq 3{\cdot}5$ sec. The corresponding resonance line width is less than 10^{-4} G. In contrast, for ice at -40°C, the resonance line width is 16 G ($T_2 \simeq 5 \times 10^{-6}$ sec), and T_1 is about 30 sec.

Consider now the hydrogen nuclear relaxation times for water adsorbed on a diamagnetic solid surface. The resonance line width in this case gives us a qualitative measure of the motional freedom enjoyed by the adsorbed molecules. We may vary the degree of surface coverage. We may also vary the adsorbing species. Estimates of binding

energy may be made from the temperature dependence of T_2.

Very often there will exist for a surface more than one way in which a molecule may be adsorbed. In general, each adsorbing phase will have its own set of characteristic relaxation times. Observation of a distribution of relaxation times T_1 and/or T_2 is indicative of multi-phase adsorption. Nuclear magnetic resonance techniques often are capable of unraveling such multi-phase behavior, and in addition yield under suitable conditions the populations of each phase.

Since the resonance signal strength is proportional to the number of contributing nuclei, the degree of surface coverage can be determined if the sample surface area is known. When it is of interest, surface coverage can be monitored as a function of time of exposure of the adsorbent to the adsorbing species.

It is instructive to consider some of the pioneering work done in this field. In their work on water molecules adsorbed on silica gel, ZIMMERMAN *et al.*[2, 3] were able to demonstrate, by nuclear resonance of the hydrogen nuclei of the adsorbed water, that for low surface coverage there exist two discrete adsorption energy sites. This fact was established by observation of two-phase behavior in the relaxation time T_2. As surface coverage is increased, a third phase, presumably due to water adsorbed on top of the first monolayer, is observed in T_2. The relative populations of the phases were determined also. Close estimates were obtained for the lifetime of a water molecule in an adsorbing phase at low coverage on silica gel, and some

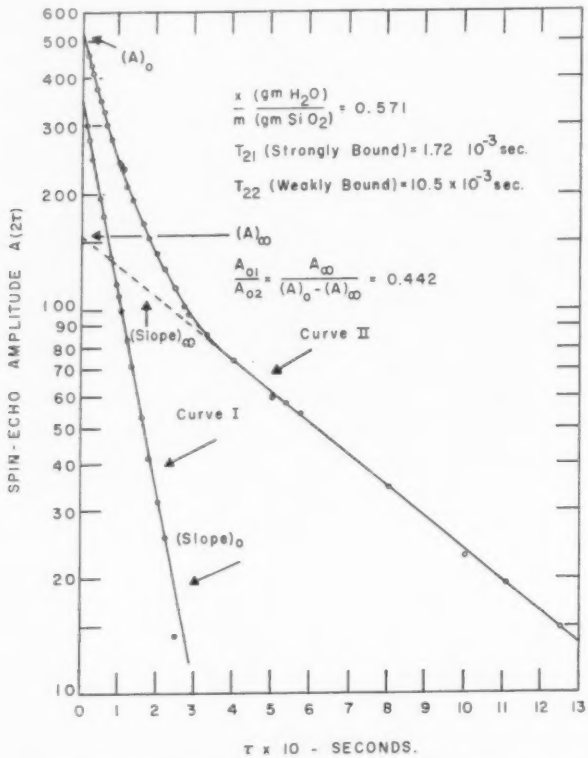


FIG. 3. Spin–spin relaxation times for H_2O adsorbed on silica gel (after ZIMMERMAN *et al.*[2]).

VOL.
14
1960

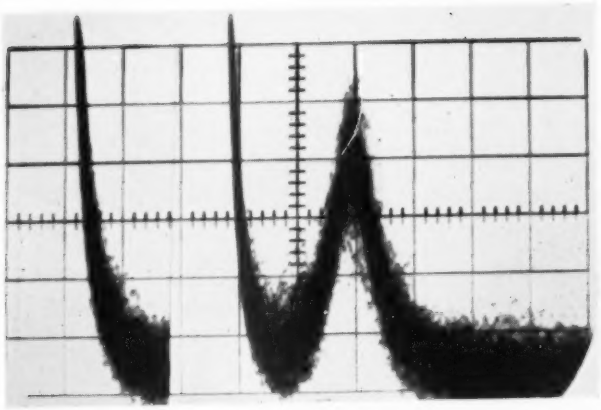


FIG. 5. Spin-echo signal from H_2O adsorbed on powdered silicon.

[facing p. 48

VOL
14
196

limits could be established on the various binding energies.

Figures 3 and 4 are reproduced from ZIMMERMAN *et al.* Fig. 3 is a typical plot for water on silica gel of spin-echo signal amplitude $A(2\tau)$ as a function of τ, the time between r.f. pulses (see Section 2). If the adsorbed system were a single phase, such a plot would yield a straight line. In

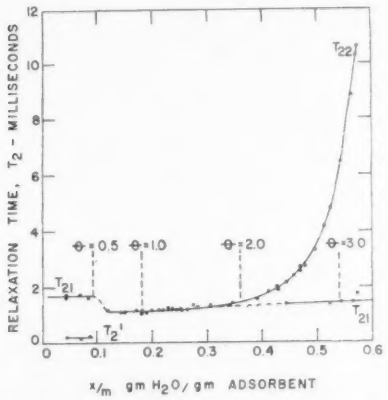


FIG. 4. Spin–spin relaxation times vs. degree of surface coverage for H_2O adsorbed on silica gel (after ZIMMERMAN *et al.*[2]).

actuality, the slope of curve (2) at large τ corresponds to one adsorbing phase, while curve (1)—which is a plot of the difference between curve (3) and the extrapolated curve (2)—corresponds to a second adsorbing phase. The ratio A_{01}/A_{02} represents the relative amounts of water in the two adsorbing phases, determined from signal amplitude measurements. Fig. 4 shows the measured relaxation times as a function of surface coverage. θ is the fraction monomolecular coverage, taking the cross-sectional area of a water molecule to be 10·6 Å^2.

In other investigations, MAYS and BRADY[5], and FUSCHILLO and ASTON[4] were able to obtain by resonance techniques a good physical picture of the behavior of water molecules adsorbed on rutile (TiO_2). FUSCHILLO and ASTON[4] also studied methane adsorbed on rutile. TANAKA[1] has reported the observation of two-phase behavior for water adsorbed on carbon.

The preceding discussion is intended to illustrate the various types of information one might hope to obtain by application of nuclear resonance techniques to nuclei of molecules adsorbed on semiconductors. The author and C. G. MCCORMICK are attempting to apply such techniques to molecules adsorbed on silicon surfaces. As was pointed out in Section 2, we are faced with a serious sensitivity problem. Whereas for silica gels and rutile one is dealing with large surface areas per unit mass (120 m^2/g to 600 m^2/g), in the case of silicon, even silicon which is crushed to a powder, one cannot hope to obtain surface areas much greater than ~5 m^2/g. Since at room temperature at least 10^{19} nuclei are required for detection of the resonance signal, relatively high surface coverages are necessary.

In our preliminary experiments, silicon was powdered in air by the technique described in Section 3. Average particle size was about 0·7μ. The powder was pressed into a porous plug of length ≈2 cm and cross-sectional area ≈3 cm^2, a nearly-saturated moist nitrogen was passed through it at a rate of about 1 cm^3/sec.

With our present sensitivity, we can detect a signal at room temperature from about one monomolecular layer of water. We expect ultimately to be able to detect about 0·1 monomolecular layer.

Figure 5 is a photograph of the spin-echo signal observed at room temperature from H_2O adsorbed on silicon after passage of moist nitrogen through the porous plug for about 1–1/2 days. The coverage in this case is several monomolecular layers. Under these conditions we find $T_1 \simeq 8{\cdot}5 \pm 1{\cdot}0$ msec and $T_2 \simeq 0{\cdot}6 \pm 0{\cdot}1$ msec. At present we are uncertain as to whether there is a multi-phase behavior in the relaxation times.

The paramagnetism associated with mechanically damaged semiconductor surfaces, discussed in Section 3, may prove to be a problem in the nuclear resonance study of species adsorbed on semiconductors. Since electron magnetic moments are of the order of one thousand times larger than nuclear moments, the former may be extremely effective in nuclear relaxation, and hence may complicate the type of nuclear resonance measurements we have considered. This problem can be evaluated more accurately at a later date when more is known about the nature of the surface paramagnetism. In particular, the effect of annealing would be of interest here.

On the other hand, it is possible that the surface

paramagnetism may be studied by means of its interaction with adsorbed nuclear magnetic species. Dynamic polarization experiments of the type first accomplished by ERB *et al.*[20] for benzene adsorbed on charred hydrocarbon surfaces may yield information on the relative motion of the paramagnetic surface centers with respect to adsorbed species, as well as information on the type of coupling (dipolar or scaler) which is predominant. If polarization can be accomplished, it could provide a way to increase sensitivity for nuclear resonance of adsorbed species by a factor of up to several hundred. We plan to attempt dynamic polarization of hydrogen nuclei in water adsorbed on silicon.

Acknowledgements—The author wishes to thank Dr. C. G. MCCORMICK and Dr. T. L. ESTLE for their participation in the experiments described and for many enlightening conversations. Discussions of the theory of surfaces with Dr. R. STRATTON have been extremely helpful to the author.

REFERENCES

1. TANAKA K. and YAMAGATA K., *Bull. Chem. Soc. Japan* **28,** 90 (1955).
2. ZIMMERMAN J. R., HOLMES B. G. and LASATER J. A., *J. Phys. Chem.* **60,** 1157 (1956).
3. ZIMMERMAN J. R. and BRITTIN W. E., *J. Phys. Chem.* **61,** 1328 (1957).
4. FUSCHILLO N. and ASTON J. G., *J. Chem. Phys.* **24,** 1277 (1956).
5. MAYS J. M. and BRADY G. W., *J. Chem. Phys.* **25,** 583 (1956).
6. FEHER G., *Phys. Rev.* **114,** 1219 (1959).
7. HANDLER P. This Conference, p. 1.
8. INGRAM D. J. E., *Spectroscopy at Radio and Microwave Frequencies.* Butterworths, London (1955).
9. ANDREW E. R., *Nuclear Magnetic Resonance.* University Press, Cambridge (1955).
10. BLEANEY B. and STEVENS K. W. H., *Reports on Progress in Physics* Vol. 16, p. 108. Physical Society, London (1953).
11. PRYCE M. H. L., *Nuovo Cim. Suppl.* **6,** 817 (1957).
12. FEHER G., *Bell. Syst. Tech. J.* **36,** 449 (1957).
13. PORTIS A. M., *Phys. Rev.* **91,** 1071 (1953).
14. HAHN E. L., *Phys. Rev.* **80,** 580 (1950).
15. CARR H. Y. and PURCELL E. M., *Phys. Rev.* **94,** 630 (1954).
16. FLETCHER R. C., YAGER W. A., PEARSON G. L., HOLDEN A. N. and MERRITT F. R., *Phys. Rev.* **94,** 1392 (1954).
17. ELLIOTT R. J., *Phys. Rev.* **96,** 266 (1954).
18. CLARKE E. N. and HOPKINS R. L., *Phys. Rev.* **91,** 1566 (1953).
19. BUCK T. M. and MCKIM F. S., *J. Electrochem. Soc.* **103,** 593 (1956).
20. ERB E., MOCHANE J. L. and UEBERSFELD J., *C. R. Acad. Sci. Paris* **246,** 1833; 2121; 3050 (1958).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 51–55. Printed in Great Britain.

CHEMICAL TECHNIQUES FOR THE INVESTIGATION OF SURFACES

R. J. KOKES

Johns Hopkins University, Department of Chemistry, Baltimore, Maryland

Abstract—A microcatalytic technique in which studies are made with very small amounts of reactant and catalyst is discussed. Applications of this technique to studies of catalytic activity, both steady and non-steady state, are reviewed and some obvious extensions are noted. Tracer investigations aimed at the study of the nature of the surface itself are also discussed.

VOL. 14 1960

IN this paper we would like to describe briefly two approaches that have been used in investigations of the nature of catalytic surfaces. Broadly speaking, these approaches could be classified as microcatalytic techniques and tracer techniques. Although the particular applications we shall discuss are of interest primarily to catalytic chemists, the techniques themselves should find applications in other studies of the nature of solid surfaces.

THE MICROCATALYTIC REACTOR

In the last decade gas chromatography and thin-film chromatography have proved to be an amazingly versatile tool in the analysis of mixtures of volatile liquids.[1–3] The operation of such devices (see Fig. 1) has been described in detail by a number of authors.[4] The mixture to be analyzed is added to a stream of gas such as helium or nitrogen. This gas is passed through a column that

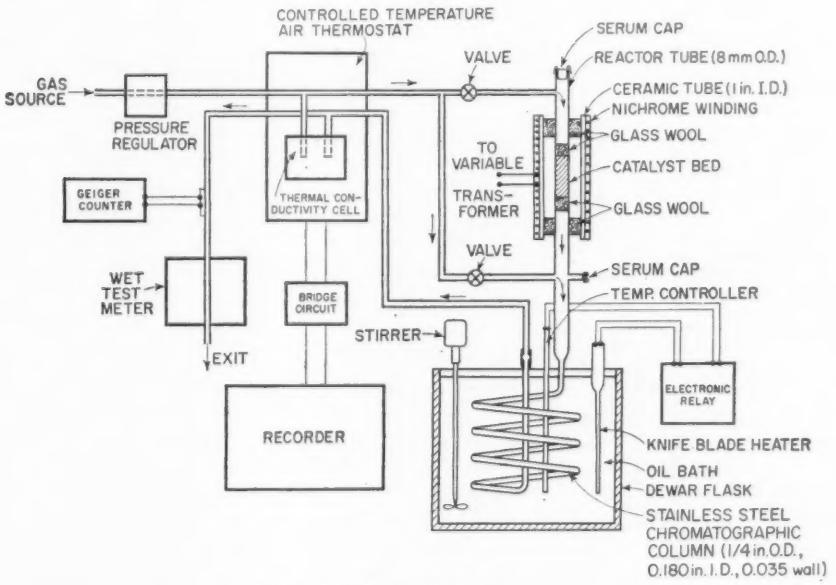


FIG. 1. Microcatalytic reactor.

contains either an adsorbent coated with a thin film of non-volatile liquid (such as paraffin oil) or an uncoated adsorbent. During the passage of the unknown mixture through the column, all the components are adsorbed by the packing. Those least strongly adsorbed appear first in the exit stream and the other components come out in an order determined by the strength of adsorption. For a given column under a specified set of conditions there will be a characteristic retention time for every pure compound. Thus, a complex mixture injected as a single slug into the inlet gas stream will emerge split into several slugs. If the choice of columns has been a wise one, each of these slugs will correspond to one of the components of the original mixture. If the thermal conductivity or some such property of the exit stream is measured with a recording potentiometer, we would obtain a curve like that shown in Fig. 2. In the ideal case,

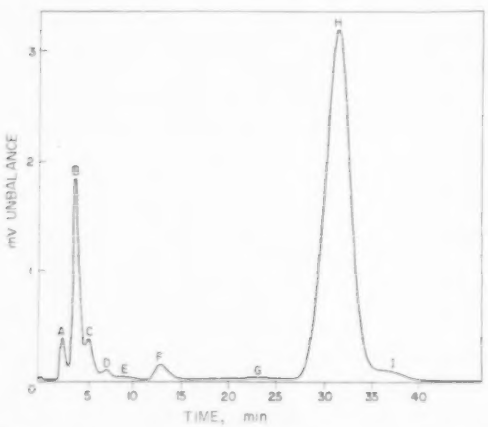


FIG. 2. Chromatogram of the products from the decomposition of 2,3-dimethylbutane over a cracking catalyst at 540°C.

each peak in this chromatogram would correspond to one component of the original mixture and the area under each peak provides a quantitative measure of the amount of each component. This relatively simple apparatus is an analytical tool of surprising range; not only can it completely separate multicomponent complex mixtures, it can even perform complete isotopic separation.[5] The sensitivity of this device is also impressive; with recently developed techniques it is possible to detect as little as 10^{-11} moles of a given component.[3]

The activity of catalysts can be measured by passing reactant gases through a tube containing the catalyst and analyzing the effluent gases in order to determine the extent of reaction. Theoretical treatment of such data is possible provided a true steady state is reached. More often than not, the steady state is not reached instantaneously, but is preceded by a gradual increase or decrease in activity with time that may by accompanied by a change in product distribution. Such behavior indicates that the nature of the surface is undergoing progressive change and that the "lined-out" activity of the catalyst is characteristic of a surface different from the initial surface. Studies of this initial non-steady state would provide valuable information on the changes that take place on the catalytic surface. Such studies can be conveniently made by attaching a microreactor unit to the top of a chromatographic column as shown in Fig. 1.[6]

In microcatalytic studies small amounts of reactant (slugs) are introduced up-stream to the reactor. This can be done by injection of the reactant into the carrier gas stream through a serum cap[6] or by a suitable bypass arrangement.[7] Typical results of such a one-shot study of the decomposition of 2, 3-dimethylbutane over a silica–alumina cracking catalyst[6] at 540°C are shown in Fig. 2. In this particular experiment, 27 μl. of reactant was injected into the stream of hydrogen carrying gas and passed through 1 cm^3 of catalyst. The peak labelled *H* corresponds to unreacted 2,3-dimethylbutane; the other peaks labeled *A*, *B*, *C*, *D*, *E*, *F*, *G*, etc., correspond respectively to: C_2H_4 and C_2H_6; C_3H_8 and C_3H_6; *i*-C_4H_{10}; *n*-C_4H_{10}, *i*-C_4H_8 and *n*-C_4H_8–1; *cis* and *trans*-C_4H_8-2; 2-methobutane and 2–2 dimethylbutane or 2-methyl–2 butane. The products formed are comparable to those reported by GREENSFELDER and VOGE[8] who also studied this particular reaction.

This rapid one-shot technique can also be adopted to tracer experiments[6] if we add a flow type Geiger counter[9] in series with the detector as shown in Fig. 1. It is well known that olefins polymerize over cracking catalysts. Fig. 3 shows the results of a run in which ethylene labeled with carbon-14 and non-radioactive propylene was injected into the microcatalytic reactor. The dotted line shows the chromatogram and the solid line

VOL. 14 1960

represents the output of a counting ratemeter. Thus, this single curve yields both the product distribution and radiocarbon assay of the products.

Figure 4 shows data for several successive runs on the micro-catalytic reactor in a study of the polymerization of propylene over a silica–alumina catalyst.[10] In this set of runs the catalyst was being poisoned by something present in the feed gas. As a result, both the over-all activity of the catalyst and the product distribution changed rapidly with time. In a steady-state flow system the changes readily detected by this slug technique would have occurred in far less than a minute. If the sensitivity available with the latest chromatography equipment has been fully exploited, it would have been possible to carry out these series of runs with reactant-catalyst contact times comparable to those that would be used in steady-state studies, but with exposure of the catalyst to reactants for periods less than a second. In such a study one could obtain the intrinsic catalytic activity of the initial catalyst, and by careful mass

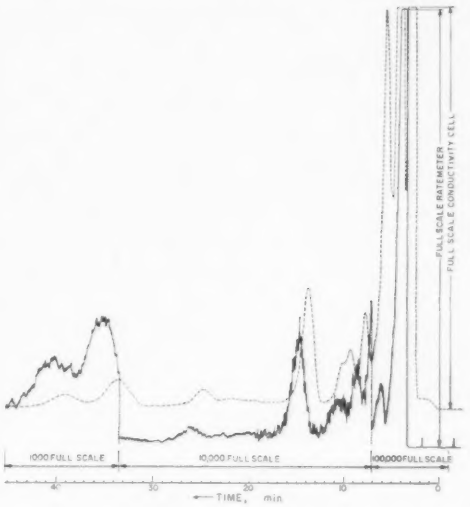


FIG. 3. Chromatogram (broken line) and radiocarbon assay (solid line) for the polymerization products of radioactive ethylene and nonradioactive propylene. Note different scales for different portions of radiocarbon assay curve.

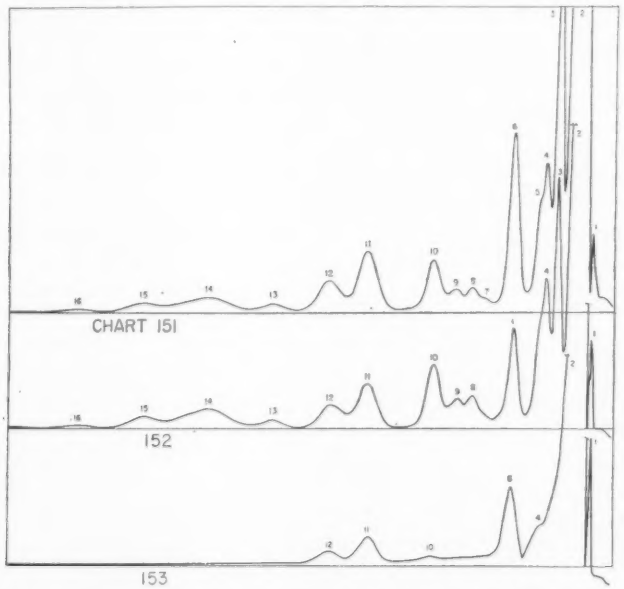


FIG. 4. Changes in selectivity and activity of cracking catalysts after three successive runs in which propylene was injected into the reactor. The polymerization products formed include: various C_4 hydrocarbons, labeled 3,4,5; C_5 hydrocarbons, labeled 6, 7, 8, 9, 10; as well as C_6 hydrocarbons, labeled 10, 11, 12, 13, 14, 15, 16. The chromatograms also show the reactant 2 together with C_1 and C_2 fragments formed by cracking.

balance, one could also gain information about changes occurring on the catalyst surface due to reaction with the feed.

Recently, studies of the Fischer–Tropsch synthesis[11] have been made in this nonsteady-state, initial stage of catalytic reaction. This reaction of carbon monoxide and hydrogen to form hydrocarbons is generally carried out in the presence of a reduced iron catalyst. It is well-known that this catalyst may take as much as 24 hr operation on stream before it reaches a steady-state. It is also well-known that a used Fischer–Tropsch catalyst contains large amounts of iron carbides and oxides; so much so, in fact, that it is questionable whether metallic iron is the catalyst. Fig. 5 shows

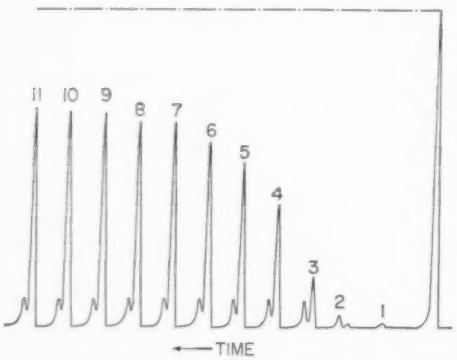


FIG. 5. Initial reactions of hydrogen and carbon monoxide (1:1 ratio) over a reduced iron catalyst at 240°C. The first peak shows the deflection obtained for the unseparated reactant gases. The peaks labeled 1, 2, etc., indicate successive runs. In each run the first peak corresponds to reactant; the second to carbon dioxide.

the results of a microcatalytic study on this system.[11] The products of the initial reaction of carbon monoxide and hydrogen on the reduced iron catalyst are almost exclusively carbon dioxide. The amount of such reaction falls with successive slugs of reactant but in each of the runs recorded in Fig. 5 no detectable amounts of hydrocarbon are formed. Mass balance suggests that the carbon dioxide is produced wholly by carbide formation. Additional experiments of this type have indicated that a carbided iron catalyst produces far more hydrocarbons in the initial stages of the reaction than the reduced iron catalysts.

The microcatalytic technique has recently been used in several catalytic studies.[12–14] This technique may be preferred over the steady-state technique when the catalyst is very susceptible to poisoning, i.e. reaction with the reactants. With this technique it is possible to keep the exposure time of the catalyst to reactant gases small by using very small slugs of gas. At the same time, since the time of contact of reactant with catalysts depends only on flow rate (ignoring volume changes), this value can be made comparable to that which would be used in a steady-state flow system. Understandably, the amount of catalyst poisoning by reactants depends on exposure time rather than contact time. Clearly, if the former is made small by the use of microcatalytic techniques, the poisoning will be minimized. The effectiveness of this technique can be readily seen by comparison of analysis for successive slugs.[12, 13] These techniques are also useful when exothermic reactions are studied on very active catalysts. In such cases the catalyst has a tendency to "run away": the heat of the reaction increases the catalyst temperatures; this increases the rate of reaction which further increases the catalyst temperature; consequently, the reaction goes nearly to completion independent of bath temperature. Since thermal effects also are minimized if the exposure time is short, the use of microcatalytic techniques greatly facilitates such studies.[12, 14]

These techniques should prove useful in studies of the catalytic activity of semiconducting oxides. In reactions such as the ammonia oxidation[15] it has been well-established that the catalyst activity is markedly influenced by the incorporation of oxygen from the reactants into the catalyst. It is, in principle, possible to add or remove traces of oxygen before a microcatalytic run and to see how the catalyst oxygen content varies after exposure to the reactants. In an extended study with very small slugs, it should be possible to study, in detail, variation of the excess oxygen with exposure time to reactants and also to determine with a high degree of precision the dependence of catalytic activity on the excess oxygen present.

TRACER TECHNIQUES

The literature dealing with tracer studies on catalysts is rather overwhelming. In this section we would like to point to three such studies which have been used to characterize the surface of a solid.

It was remarked earlier that carbided iron catalysts are more effective in the Fischer–Tropsch synthesis than reduced iron. Since carbides can be hydrogenated to methane, it was once believed that the carbide functioned as an intermediate in the higher hydrocarbon synthesis. In an effort to test this hypothesis KUMMER and EMMETT[16] prepared an iron catalyst carbided with radioactive carbon-14. This catalyst was then used for hydrocarbon synthesis with nonradioactive reactants. The radioactivity of the products was such that not more than 15 per cent of them could have been synthesized *via* the carbide.

All heterogeneous catalytic reactions are assumed to proceed subsequent to chemical adsorption on a solid. In some cases, only a small fraction of the available chemisorption sites are active for catalysis. Nevertheless, it would seem that all catalytic reactions should be accompanied by chemisorption. One exception to this seemed to be the catalytic cracking of alkanes on silica–alumina catalysts. Standard measurements[17] revealed that although physical adsorption of alkanes did occur, chemisorption was undetectable and must be less than 10^{-3} cm^3 STP/g on a catalyst which would form a monolayer containing 40 cm^3 STP/g. Later studies utilized radioactive alkanes.[18] With these methods it was possible in principle to detect chemisorptions of 10^{-8} cm^3 STP/g in the presence of physically adsorbed gases. Such studies revealed that the amount of strongly held alkane is about 0·0005 cm^3/g. Presumably this relatively small number of active sites accounts for the activity of silica–alumina catalysts.

The beautiful studies carried out by WINTERS and co-workers[19] which deal with O^{18} exchange with solids, have added a great deal to our knowledge of oxide surfaces. The conclusions drawn from the work are many but one of the more surprising is the fact that gas phase oxygen will exchange with solids such as Fe_2O_3 at temperatures as low as 300°C. Moreover, the amount of readily exchangeable oxygen agrees quite well with the amount on the surface as estimated by the BET method. All aspects of these researches lead to the picture of a more mobile surface than was hitherto conceived.

Acknowledgement—The author would like to thank Professor P. H. EMMETT for permission to discuss herein unpublished data obtained in his laboratory.

REFERENCES

1. JAMES A. T. and MARTIN A., *Biochem. J.* **52,** 242, (1952).
2. LICHTENFELS D. H., FLECK D. H. and BURROW F. H., *Analyt. Chem.* **27,** 1510 (1955).
3. PITKETHLY R. C., *Analyt. Chem.* **30,** 1309 (1958).
4. Cf. *Vapor Phase Chromatography* (edited by DESTY P. H.). Academic Press, New York (1957).
5. OHKOSKI S., TENMA S. and KWAN T., *Bull. Chem. Soc. Japan* **31,** 772 (1958).
6. KOKES R. J., TOBIN H. Jr. and EMMETT P. H., *J. Amer. Chem. Soc.* **77,** 5860 (1955).
7. HALL W. K. and EMMETT P. H., *J. Amer. Chem. Soc.* **79,** 2091 (1957).
8. GREENSFELDER B. S. and VOGE H. H., *Industr. Engng. Chem.* **41,** 2573 (1949).
9. KUMMER J. T., *Nucleonics* **3,** 27 (1948).
10. TOBIN H. and EMMETT P. H., unpublished data.
11. DE MOURGES L. and EMMETT P. H., unpublished data.
12. HALL W. K. and EMMETT P. H., *J. Phys. Chem.* **63,** 1102 (1959).
13. PASS G. and EMMETT P. H., unpublished data.
14. KOKES R. J. and EMMETT P. H., *J. Amer. Chem. Soc.*, to be published.
15. KRAUSS W., *Z. Electrochem.* **53,** 320 (1948).
16. KUMMER J. T. and EMMETT P. H., *J. Amer. Chem. Soc.* **70,** 3632 (1948).
17. ZABOR R. and EMMETT P. H., *J. Amer. Chem. Soc.* **73,** 5639 (1951).
18. MACIVER D. S., EMMETT P. H. and FRANK H., *J. Phys. Chem.* **62,** 935 (1958).
19. WINTERS E. R. S., *Chemisorption* (edited by GARNER W. E.) Chap. 4–2. Butterworths, London (1957).

VOL. 14 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 56–59. Printed in Great Britain.

INFRARED METHODS APPLIED TO SURFACE PHENOMENA

R. P. EISCHENS

Texaco Research Center, Beacon, New York

TECHNIQUES for observation of infrared spectra of molecules chemisorbed on the surfaces of solids have proven to be of major importance in studies of catalysis. This paper is not intended to be a review of the infrared study of chemisorbed molecules since a comprehensive review of this field has been published recently.[1] Instead I will outline the experimental methods, with examples of the type of results obtained, and then consider how these methods might be applied to studies of the surface chemistry of germanium and silicon.

Infrared studies of chemisorbed molecules can be carried out most efficiently in an *in situ* cell which makes it possible to reduce the sample, treat it with gases, and observe spectra of the chemisorbed species in a single apparatus. A schematic diagram of a suitable cell is shown in Fig. 1. The cell is made by sealing infrared transparent salt windows to the ends of a pyrex cylinder. It is heated by a winding of nichrome wire. The powdered sample is supported on a salt plate which is held in place by a tripod.

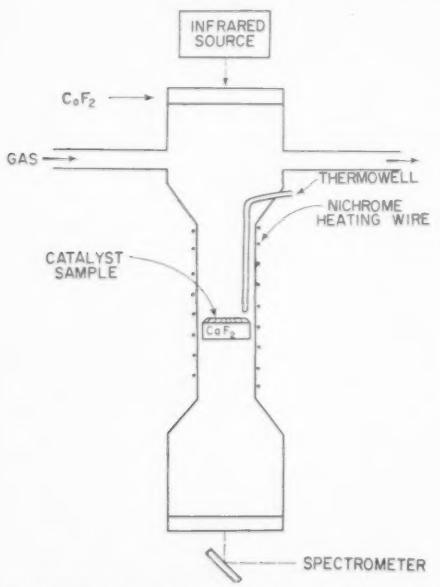


FIG. 1. Cell for infrared study of chemisorbed molecules.

Most of our infrared studies have been concerned with molecules chemisorbed on small metal particles which are dispersed on silica or alumina. The carrier is treated with a solution of a salt of the active metal, partially dried, placed in the cell, further dried and reduced.

It is important to understand the nature of this type of sample which is common in catalysis work but which is not usually encountered by workers interested in germanium or silicon. Fig. 2 is an electron micrograph of a sample of platinum (9 wt. per cent) on silica. The platinum exists as discrete particles (one is encircled) and is not spread over the silicon surface as a monomolecular layer. The platinum particle diameters fall in the range of 40–100 Å. The infrared beam goes through a large number of particles as it is transmitted through the sample. This transmission through a large number (as many as 1000) monolayers greatly simplifies the problem of obtaining detectable infrared absorption bands.

Figure 3 shows the spectra of carbon monoxide chemisorbed on silica-supported platinum and on silica-supported palladium. The spectrum of carbon monoxide chemisorbed on silica-supported platinum has a strong band of 2070 cm^{-1}. This is close to the frequency of the carbon–oxygen stretching vibration of metal carbonyls where the CO group is bonded to a single metal atom through the carbon. Thus on the basis of analogy with the spectra of metal carbonyls it was concluded that

VOL.
14
1960

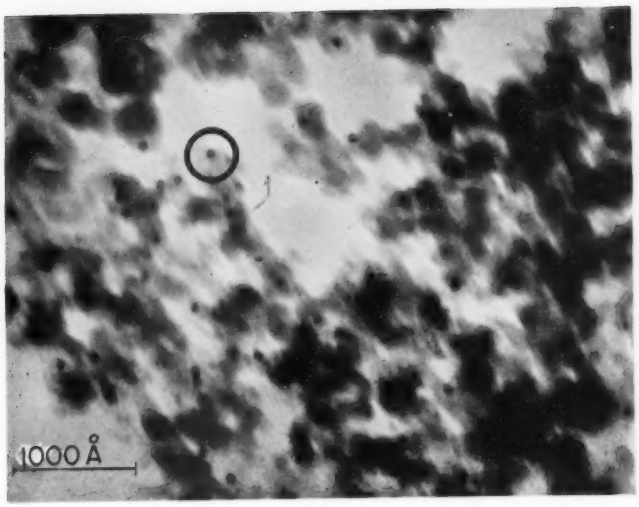


FIG. 2. Electron micrograph of silica-supported platinum.

[*facing p.* 56

VOL
14
196

VOL. 14 1960

the carbon monoxide chemisorbed on platinum is bonded through the carbon to one platinum atom as Pt–C≡O. The spectrum of carbon monoxide on palladium has strong bands at 1820 and 1925 cm^{-1} and a weak band at 2070 cm^{-1}. The latter was attributed to the linearly chemisorbed carbon monoxide as in the case of platinum. The bands

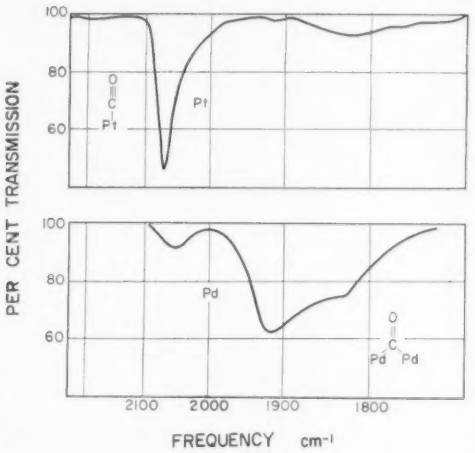


FIG. 3. Spectra of carbon monoxide chemisorbed on silica-supported platinum and on silica-supported palladium.

in the 1820–1925 cm^{-1} region are similar to those found in the spectra of carbonyls, such as $Fe_2(CO)_9$, which are due to CO groups bridged between two metal atoms. Again on the basis of analogy it was concluded that the 1820–1925 cm^{-1} bands represent the bridged type of chemisorbed carbon monoxide,

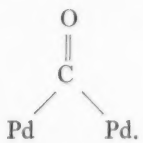


Figure 4 shows the spectrum of carbon monoxide chemisorbed on evaporated platinum films which had been deposited on the faces of sodium chloride windows. The spectrum was observed after passing the beam through a series of four films. A scale expander was used at fivefold magnification. The scale expander is an important recent development, available commercially, which in effect makes it possible to increase the sensitivity of conventional infrared spectrophotometers by as much as twenty-fold.[2]

In Fig. 4 the value of I/I_0 is 0·89 where I is the intensity transmitted at 2050 cm^{-1} after chemisorption and I_0 is the intensity transmitted prior to chemisorption. The I/I_0 value is related to c, the number of molecules per cm^2 of beam cross

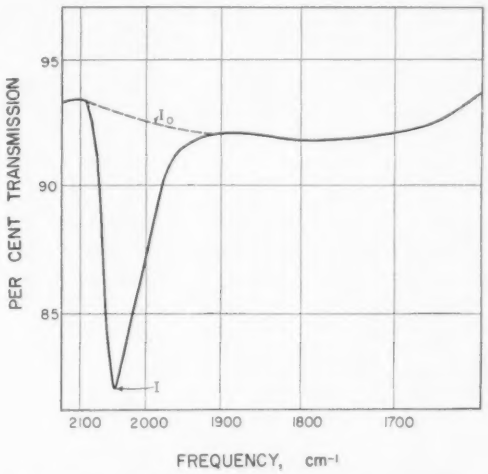


FIG. 4. Spectrum of carbon monoxide chemisorbed on evaporated platinum films.

section, by the simple absorptivity law $I/I_0 = e^{-kc}$. The extinction coefficient, k, can not be specified exactly for chemisorbed carbon monoxide because it varies with the amount of carbon monoxide which is on the surface. However, a value of $k = 300 \times 10^{-20}$ cm^2/molecule can reasonably be used to calculate an approximate value for c. When this is done, c is found to be 40×10^{15} molecules/cm^2. Since roughly 1×10^{15} molecules of carbon monoxide can be chemisorbed on one cm^2 of platinum surface, the measured I/I_0 value shows that the infrared beam has passed through the equivalent of forty monolayers of chemisorbed carbon monoxide (Fig. 4). Thus each platinum film chemisorbed the equivalent of ten monolayers of carbon monoxide. This is reasonable since films of this type are porous and can have surface areas which are greater than the geometric area.

As a final point concerning the evaporated film work it would be misleading not to mention that the platinum–carbon monoxide system is exceptionally well suited for infrared studies. This is

because chemisorbed carbon monoxide has a high extinction coefficient and because carbon monoxide is able to displace impurities such as oxygen from the platinum surface.

Without going into detail, mention should be made of infrared studies of molecules on surfaces by reflection techniques. PICKERING and ECKSTROM used 15–35 reflections from evaporated nickel surfaces to study chemisorbed carbon monoxide,[3] FRANCIS and ELLISON studied fatty acid soaps on polished metal with an optical system involving four reflections,[4] and GARDNER and PETRUCCI formed a light pipe by evaporating nickel onto the inner walls of a glass cylinder and then studied the spectra of carbon monoxide chemisorbed on the nickel.[5] It is unlikely that any of these three techniques would be applicable to studies on germanium or silicon because they demand surfaces of high reflectivity.

The dispersed metal technique which was described first is ideally suited for study of systems of interest to catalysis but it is not well suited for studies of germanium and silicon surfaces. However, it is possible to prepare samples of dispersed germanium by thermal decomposition of germane. This was done by TAMARU and BOUDART[6, 7] who prepared germanium with a surface area of 2·6 m^2/g by decomposing germanium glass fibers. It seems likely that a higher area germanium could be prepared by decomposition of germane on the small particles of silica which is used as a carrier in the dispersed metal samples. However, a germanium area as low as 2·6 m^2/g is sufficient to give fifty layers of chemisorbed gas in the beam path if it is assumed that 0·002 g/cm^2 (of beam cross section) could be tolerated. This would be equivalent to the amount of metal which is used in the dispersed metal samples.

The dispersed germanium samples might be satisfactory for some problems of special interest in germanium surface chemistry. For example information on whether the surface of the germanium is contaminated by oxide could be obtained by observing the spectrum of chemisorbed hydrogen. Hydrogen chemisorbed on an oxygen atom would produce hydroxyl bands while hydrogen chemisorbed on germanium would produce infrared bands analagous to those of germane. This type of sample would not be satisfactory for observing the oxygen on germanium directly because the germanium–oxygen bands would be obscured by silicon–oxygen bands of the carrier. Evaporated silicon or germanium films would not suffer from this disadvantage.

HASS and SALZBERG[8] observed silicon–oxygen bands with I/I_0 values (at 1000 cm^{-1}) of 0·78 after passing through silicon monoxide films which were 1250 Å thick. The work of KAISER *et al.*[9] indicates that it is reasonable to assume that germanium-oxide bands would have about the same extinction coefficient as silicon–oxygen bands. If one assumes that the 1250 Å silicon monoxide films are equivalent to 200–300 monolayers of chemisorbed oxygen it is evident that techniques similar to those used to obtain Fig. 4 could be expected to detect bands due to oxygen chemisorbed on silicon or germanium.

A similar conclusion is indicated for attempts to observe chemisorbed oxygen on germanium or silicon disks. Because of reflection losses of about 50 per cent for each disk, probably not more than three or four disks could be used in series. There is an added complication due to molecular orientation which might be important in work with disks if these surfaces approach atomic smoothness. Molecular vibrations which produce dipole changes normal to the plane of the surface will not absorb radiation which is normal to the surface plane. This effect is not expected to be of importance in the work with dispersed metal samples because molecules chemisorbed on the small metal particles are randomly oriented with respect to the beam. Attempts to detect an orientation effect for carbon monoxide on evaporated platinum films (by changing the angle of incidence) were not successful probably because there is also no significant orientation in this case.[10]

In summary it appears that there are no insurmountable obstacles to applying the infrared techniques to studies of molecules chemisorbed on germanium or silicon. There is reason to believe that in a few specific cases information of interest could be obtained using commercially available infrared equipment fitted with thermocouple detectors. However, it is also evident that the high surface to volume ratio of the samples which are of interest to catalysis is an important advantage in the application of the infrared method to problems associated with this field.

REFERENCES

1. EISCHENS R. P. and PLISKIN W. A., *Advanc. Catalys.* **10,** 2 (1958).
2. COATES V. J. and ANACREON R., reported at *Conference on Applied Spectroscopy, Pittsburgh,* 1957.
3. PICKERING H. L. and ECKSTROM H. C., *J. Phys. Chem.* **63,** 512 (1959).
4. FRANCES S. A. and ELLISON A. H., *J. Opt. Soc. Amer.* **49,** 131 (1959).
5. GARDNER R. A., Thesis, Western Reserve University (1959).
6. TAMARU K., *J. Phys. Chem.* **61,** 647 (1957).
7. TAMARU K. and BOUDART M., *Advanc. Catalys.* **9,** 699 (1957).
8. HASS G. and SALZBERG C. D., *J. Opt. Soc. Amer.* **44,** 181 (1954).
9. KAISER W., KECK P. H. and LANGE C. F., *Phys. Rev.* **101,** 1264 (1956).
10. EISCHENS R. P. and PLISKIN W. A., unpublished results.

VOL.
14
1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 60–71. Printed in Great Britain.

SEMICONDUCTOR SURFACE PROPERTIES DEDUCED FROM FREE CARRIER ABSORPTION AND REFLECTION OF INFRARED RADIATION

N. J. HARRICK

Philips Laboratories, Irvington-on-Hudson, New York

Abstract—The use of infrared radiation to study semiconductor bulk properties through the absorption by the free carriers is reviewed. Numerous applications of this technique and variations thereof are then discussed whereby properties of semiconductor surfaces may be determined. It is shown that in many instances the techniques described here have distinct advantages over other techniques presently in use.

INTRODUCTION

SEMICONDUCTORS are highly transparent to infrared radiation above the lattice absorption edge. Any absorption can be accounted for by the action of the free carriers although the absorption coefficients are small.[1] Injected carriers are found to absorb radiation to the same degree as equilibrium carriers. Thus a significant modulation of infrared radiation can be brought about by the injection of added carriers.[2] By using a narrow beam of infrared radiation, the modulation of the transmitted beam has been used to probe out carrier distributions in the semiconductor bulk for injection and extraction.[3] This technique is finding wide application in checking semiconductor properties. It has been used successfully to check transport[3] and junction theories[4] and to determine sample characteristics such as lifetime,[5,3] type and homogeneity.[6]

Surface properties can also be deduced from measurements utilizing infrared radiation to probe out carrier densities. Basically, four approaches may be used to determine surface characteristics; these are:

(a) Surface properties deduced from changes in bulk carrier densities or distributions.
(b) Study of the spectral characteristics of carriers in the semiconductor space charge region.
(c) Use of infrared radiation to monitor the carrier density in the semiconductor space charge region.
(d) Surface properties deduced from reflectivity measurements.

As in the bulk measurements, the surface measurements are often made easier by the possibility of changing the carrier density. In all the bulk measurements made up to date, the infrared radiation served to probe the local conductivity which in many cases could not be done electrically. For the surface, it is proposed to determine the optical properties of the carriers as well as to make conductivity measurements. The merits of the various techniques suggested will be discussed.

EXPERIMENTAL TECHNIQUE

(a) *Free carrier absorption*

The technique for the use of infrared radiation to measure local added carrier densities is shown schematically in Fig. 1. An infrared source is focussed onto a narrow slit or pinhole. The sample is placed as near as possible to the slit to maintain a narrow beam. An infrared detector is lined up with the beam which emerges from the sample. Filters are placed in the beam preceding the sample to define the wavelength range of interest; e.g. a combination of quartz and Ge filters is useful for germanium giving a band width of 2–3·5μ. The sample and detector are mounted on separate micromanipulators which permits sample position to be determined precisely and also permits the detector to be aligned critically with the beam. A fair degree of collimation is required, although the high index of refraction of the semiconductor is a

VOL. 14 1960

big help in constraining the beam to a narrow path within the semiconductor. The quantities of interest at each position of the sample are the total transmitted radiation, I_T, prior to injection or extraction and the change in transmission, ΔI, due to injection or extraction. It is found convenient to make the display of the signal on an oscilloscope. The value of I_T is obtained by observing the change indicated by the detector when the beam is obstructed by means of a shutter. ΔI, which is related to the local excess or deficit of carriers introduced in the sample, may be induced by a passage of current through the sample, by illumination of the sample, by applying an external electric or magnetic field, by a change of gaseous ambient or by other means. The measurements are simplified considerably through the use of the ratio $\Delta I/I_T$, since no absolute measurements of intensity are then required. Because the absorption coefficients of the holes and the electrons are small, a high degree of amplification is required. This puts stringent requirements on the stability of the infrared source especially for fast systems. The large thermal capacity of a globar source is a help in the regulation problem although it is by itself insufficient. Some of the work proposed here would require the use of a fast detector which would enable one to study effects connected with the fast surface states as well as high drift field effects. We have used thermopiles and lead sulphide cells as radiation detectors. Faster detectors can be made.[7]

The course of an infrared beam through a plane parallel germanium slab is shown in Fig. 2 for oblique incidence. The various components can be calculated from simple geometrical considerations. The first two transmitted components are given by

$$I_{T1} = (1-R)^2 I_0 \exp(-k_0 d)$$

and

$$I_{T2} = (1-R)^2 R^2 I_0 \exp(-3k_0 d),$$

where R is the reflectivity of the semiconductor surface, I_0 the incident radiation intensity, k_0 the extinction coefficient, and d the length of the path for one passage through the sample. For low doping (small k_0) in germanium, an error as high as 20 per cent can be encountered by neglecting the higher order components. If the higher order components are important, it is easier to eliminate them than to take them into account mathematically. By using oblique incidence as shown in Fig. 2,

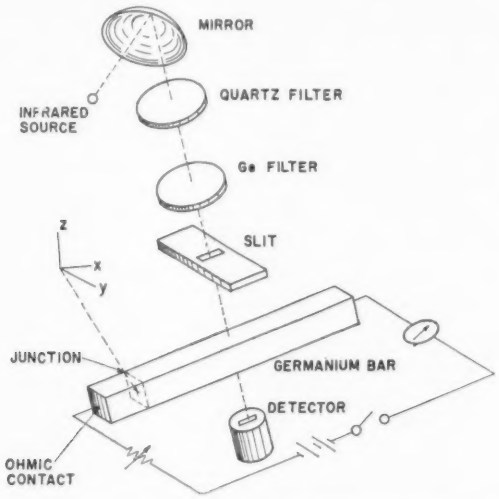


FIG. 1. Schematic diagram of the infrared set-up used to probe out carrier disturbances in semiconductors.

they can be eliminated. The higher order components can be deflected further with the aid of a small germanium wedge. It is also evident that higher sensitivity can be gained by using one of the higher order components.

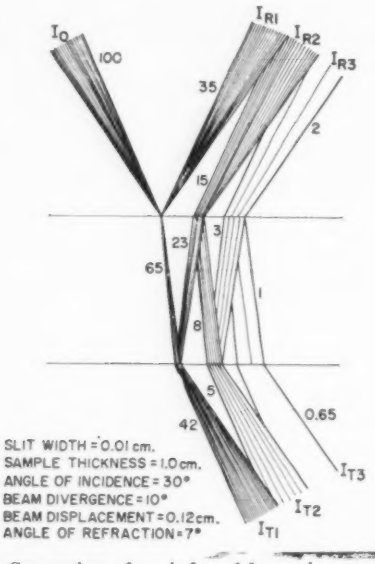


FIG. 2. Separation of an infrared beam into numerous components due to multiple reflections in a semiconductor. The numbers indicate the percentage of the original beam found in the various components for germanium.

When the carrier density is disturbed, I_{T1} can be written as

$$I_T \pm \Delta I = I_0(1-R)^2 \exp[\mp(k_0+k)d] = I_T \exp(\mp kd).$$

This is so only if the reflectivity remains unchanged with injection or extraction. The assumption that the reflectivity remains unchanged with injection is a very good one and will be discussed in more detail.

The absorption coefficients are actually wavelength dependent. By using a narrow band of radiation or by working in a region where the absorption coefficients do not change rapidly with wavelength (e.g. 2 to 5μ for germanium), an average constant may be assumed for the absorption coefficient.

If the above considerations are satisfied, the observed signal can be written as

$$\pm \Delta I/I_T = \exp[\mp(k_h\Delta p + k_e\Delta n)d] - 1,$$

where k_h and k_e are the absorption coefficients for holes and electrons, respectively. This expression reduces to

$$\Delta I/I_T \sim \pm(k_h\Delta p + k_e\Delta n)d,$$

when the exponent is much less than unity. This condition is generally met as the k's are of the order of 10^{-16} for each free carrier. For heavy injection the exponential must be used. For the neutral bulk, where it can be assumed that $\Delta p = \Delta n$, a constant proportionality exists between $\Delta I/I_T$ and the change in carrier density. Thus strict proportionality exists between Δp and $\Delta I/I_T$ if the following conditions are met:

(a) Multiply reflected beams can be neglected or eliminated.
(b) Change of reflectivity with added carrier density can be neglected.
(c) Variation of the wavelength dependence of the absorption coefficient can be neglected.
(d) Added carrier density is small.

Any non-linearity of the detector must, of course, be taken into account. Using the bandwidth 2–3½μ from a globar heated to a brightness temperature of about 1100°C, we have determined the constant relating the change in carrier density to the relative change in intensity of the infrared radiation and have found the following relationship,

$$\Delta p = \frac{5\times 10^{15}}{d}\frac{\Delta I}{I_T}.$$

The proportionality constant can be determined in one of the following ways:

(a) By a comparison of the relative optical absorption increase with electrical conductivity measurements for the same level of injection.
(b) By comparing the maximum relative increase of transmission for a near intrinsic sample under the condition of field-aided extraction to the calculated maximum permissible extraction at the same temperature.
(c) By determining separately the absorption coefficients from sample transmission measurements.

This technique can be used to probe out changes in carrier densities in the surface barrier by ininternally reflecting the infrared radiation. The added carrier densities and the fractional change

in absorption, if no changes in the carrier density in the neutral bulk occur, are related through

$$\pm \frac{\Delta I}{I_T} = 1 - R^n \exp\left[\mp(k_h \Delta P + k_e \Delta N)\frac{2n}{\cos\theta}\right].$$

Here ΔP and ΔN represent the change in the total number of excess holes and electrons per cm^2, respectively, θ the angle of incidence and n the number of reflections. It is convenient to work at angles greater than the critical angle when R becomes unity. As ΔP is not in general equal to ΔN, the individual absorption coefficients must be known.

VOL. 14 1960

(2) *Reflection by free carriers*

Although it was assumed in the previous section that the change in the reflectivity due to carrier injection is small and could be neglected in making absorption measurements, it is just this change which needs to be measured in order to determine properties of the surface from reflectivity measurements. To make such measurements, the beam is reflected internally or externally from a semiconductor surface and the quantity $\Delta I/I_R$ should then be related to the change in carrier density. The theory involved is an application of Fresnel's equations to a semiconductor. The free carriers can make a small contribution to the reflectivity through their effect on the imaginary part of the absorption coefficient. Generally this is a very small effect and the experiment becomes almost impossible if the reflectivity changes need to be separated from the absorption changes which are some 10^5 times greater, although of opposite sign.

(3) *Techniques for creating carrier disturbances*

All of the studies outlined here depend on the possibility of altering the carrier density within the semiconductor surface or bulk, at least they are made easier because this can be done. The changes in the bulk are caused by injection or extraction which actually changes the product of the carrier densities even though the temperature remains fixed. This change can be brought about through the use of junctions, injecting contacts, exciting radiation or deflection of equilibrium or added carriers through the use of crossed electric and magnetic fields. This latter technique has been highly neglected even though it appears to be very important. By using crossed electric and magnetic fields to deflect injected[8] equilibrium carriers or to deflect carriers from a surface of high generation rate to one of low generation rate,[9] added concentrations of 10^{18}/cm^3 can probably be developed in thin regions near a surface. More will be said about this when specific applications arise.

The surface carrier density can be changed either through a change in the carrier composition in the surface barrier without a change in the product of the carrier densities or through a change in the *pn* product. A change in barrier height alone can be brought about by a change in the chemistry of the surface or through a field effect type of experiment. Any change in the *pn* product in the neighboring bulk results in a change of the *pn* product in the surface, hence any of the techniques used to change the carrier density in the bulk also apply to the surface.

APPLICATIONS TO SURFACE STUDIES

(1) *Surface properties deduced from bulk carrier distributions*

(a) *Surface recombination velocity.* Perhaps the most fundamental and direct measurement of surface recombination velocity can be obtained from a measure of excess or deficit transverse carrier distribution in a semiconductor bulk resulting from injection or extraction of carriers.[3] The technique used to obtain results here is identical to that described in connection with Fig. 1 from determinations of transverse distributions. Such measurements are shown in Fig. 3 for both injection and extraction of carriers into a parallelopiped for the case that the opposing sides have unequal values of s. If the two opposing sides have equal values of s, its value is given by

$$s = \frac{D}{Y}\left(\frac{(1-\zeta_Y{}^2)^{1/2}}{\zeta_Y}\right)\cos^{-1}\zeta_Y,$$

where ζ_Y is the ratio of the value of $\Delta I/I_T$ at the surface ($y = Y$) to its value in the center of the crystal. The sample must be large enough so that the distributions can be determined precisely. Homogeneity in sample doping is required otherwise the equations involved become complicated.

A rather useful application of this technique has been made in the measurement of the surface recombination velocity for a semiconductor before

and after a metal was brought into contact with the surface to determine what effect the metal may have on the surface recombination velocity.[10] To measure conductivity electrically in this case was, of course, out of the question.

solution of the equations involved would have to be considered, which would result in complicated expressions relating surface recombination velocity and added carrier density. Crossed electric and magnetic fields can, however, be used to control the

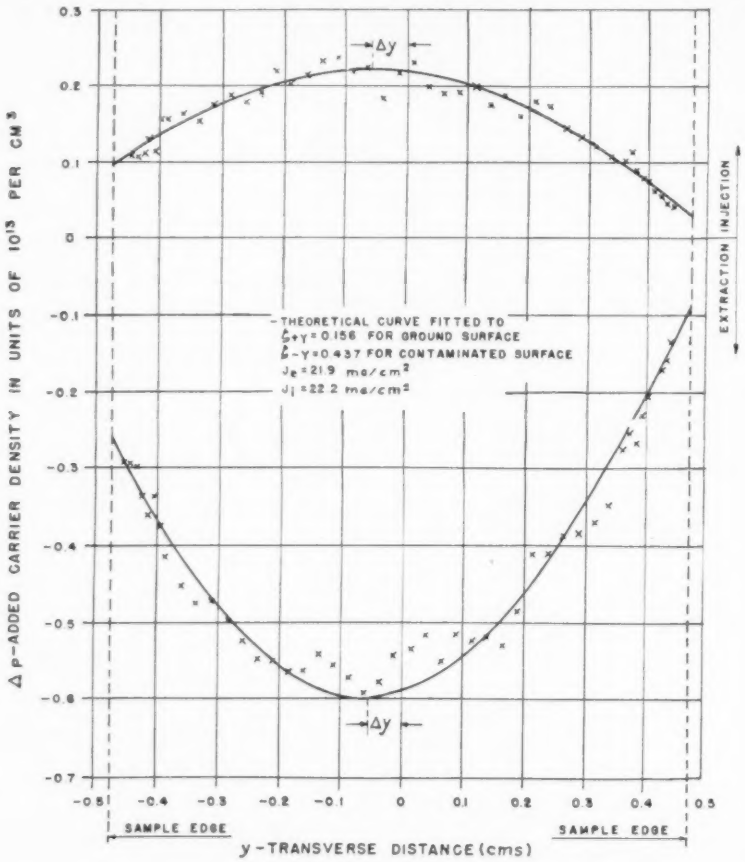

Δy
THEORETICAL CURVE FITTED TO
S+γ = 0.156 FOR GROUND SURFACE
S-γ = 0.437 FOR CONTAMINATED SURFACE
Je = 21.9 ma/cm²
Ji = 22.2 ma/cm²
EXTRACTION INJECTION
Δ p-ADDED CARRIER DENSITY IN UNITS OF 10¹³ PER CM³
SAMPLE EDGE
SAMPLE EDGE
y-TRANSVERSE DISTANCE (cms)

FIG. 3. The transverse added carrier distribution for injection and extraction in a sample whose opposing surfaces have different values of recombination velocity. Both the added and deficit carrier distributions lead to the same value of surface recombination velocity.

Another useful application of this technique might be to study surface recombination velocity versus carrier concentration. The carrier concentration cannot conveniently be controlled over a wide range through injection by a junction or light for this purpose. High concentrations could be obtained only by making measurements near the plane of injection, thus the "higher modes" in the concentration of carriers in a uniform manner and the densities can be calculated. Carriers can be deflected from a diamond polished surface. If the normal generation rate is inadequate, photo generation can be used to enhance it. Measurements of the carrier distribution can then be made in a direction parallel to the direction of the magnetic field. The infrared will pass through a region of

injection and a region of extraction; i.e. the infrared will see both of the curves of Fig. 3. The contribution of the extraction signal can be determined or can be neglected as it will be small.

Infrared radiation can be used in any other experiment where surface recombination velocity is determined from conductivity measurements. Its uniqueness lies in its use to probe out local conductivity as discussed above.

(b) *Injection or extraction properties of a surface.* When a transient or steady state current is passed through a surface, injection or extraction effects may be observed in the semiconductor bulk. From these observations it is possible to determine the surface type, the magnitude of the surface potential and the properties of the surface states. The results will vary according to the type of contact that is made to the surface. The contact may be metallic (Fig. 4(a)) or capacitive (Fig. 4(b)).

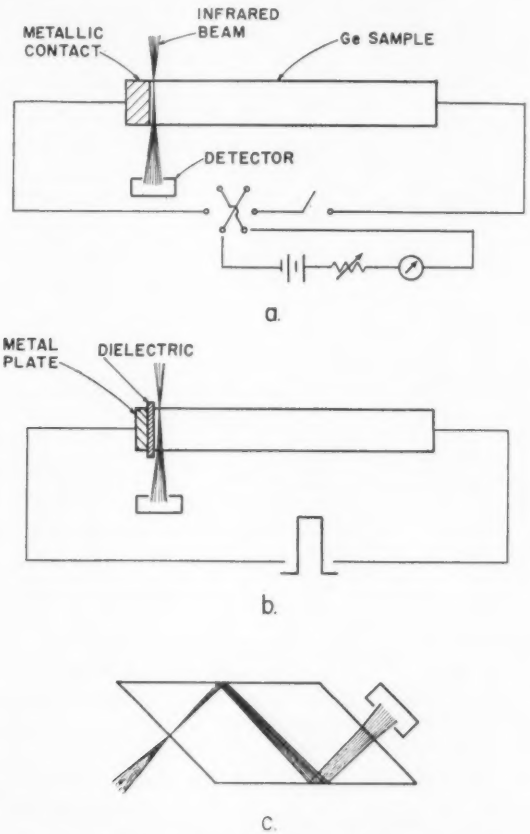


FIG. 4. Schematic diagrams of various techniques used for obtaining information about the semiconductor surface through added carrier absorption. (a) Metal to semiconductor contact; (b) capacitive contact; (c) probing the space charge region.

(i) *Metallic contact.* For the case that a metal makes contact to the semiconductor, we are dealing with the old problem of metal to semiconductor contacts. It has been convincingly shown that by examining the injection–extraction effects in the semiconductor bulk as a function of current through the use of infrared radiation,

VOL. 14 1960

much more can be determined about the nature of a contact and the surface than through an examination of the current–voltage characteristics. Thus this is a powerful technique to study metal to semiconductor contacts[11, 10] as is illustrated by the current versus added carrier density characteristics in Fig. 5 and the corresponding

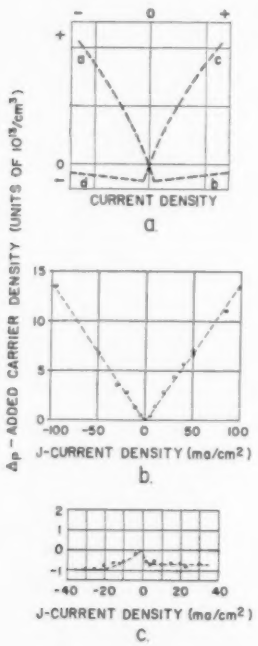


FIG. 5. Current versus added carrier density characteristics for p–n junctions and metal-to-semiconductor contacts. (a) Characteristics for grown and alloyed junctions and rectifying metal-to-semiconductor contacts; (b) injection for either direction of current flow; (c) extraction for either direction of current flow. The latter two curves may be observed for a contact where the metal and the semiconductor are separated by a thin insulating layer.

current–voltage characteristics in Fig. 6. Normal rectifying contacts can be investigated through an examination of either the current versus added carrier density characteristics (Fig. 5(a)) or the current voltage (Fig. 6(a)) characteristics, although the former has some advantages.[11] Other contacts such as those exhibiting injection or extraction for either direction of current flow, however, may exhibit large effects in their current versus added carrier density characteristics (Figs. 5(b) and 5(c)) and small or no effects in their current–voltage characteristics (Figs. 6(b) and 6(c)). It was found that an insulating film between the metal and the semiconductor plays an important role in these unusual characteristics (Figs. 5(b) and 5(c)) through a field effect and a voltage controlled injection.[11] In the absence of the insulating layer, the semiconductor surface, which probably remains unaltered when the metal is brought into contact, is the controlling factor at the metal to semiconductor contact. Such contacts exhibit J–Δp characteristics quite like those for grown or alloyed junctions. The characteristics of such contacts can be described by treating the semiconductor surface–semiconductor bulk structure in a manner analogous to a p–n junction.[10] Measurements of the current versus added carrier density characteristics give information on the surface type, magnitude of the surface potential and properties of the surface states.

(ii) *Capacitive contact.* The surface injection experiment[12] (Fig. 4(b)) which is just a high frequency field effect experiment, can also be performed using infrared radiation to detect the excess or deficit of carriers in the semiconductor bulk. For this experiment it is necessary to have a detection scheme which would respond in a time of not more than the carrier lifetime. For the examination of the fast states, even a faster detector is necessary. It has already been shown that this experiment may be used to determine surface type and surface potential as well as to study surface states. It may also be useful in determining surface mobilities. This technique for studying surfaces has the advantage that no assumption need be made concerning surface mobilities as is required in the field effect experiment.

(c) *Measurement of surface barrier height.* Barrier heights under metal contacts can be determined precisely from an experiment similar to that described by JOHNSON[13] for a free surface. In the present experiment, carriers are injected optically and simultaneous measurements are made of their density under the contact and the floating potential of the contact. These experimental results can be used to identify the surface barrier height. Fig. 7 shows such measurements for silver and aluminum evaporated on cleaved, near intrinsic germanium samples. The surface was found to be strongly p-type and the barrier height was independent of which metal made contact to the

VOL. 14 1960

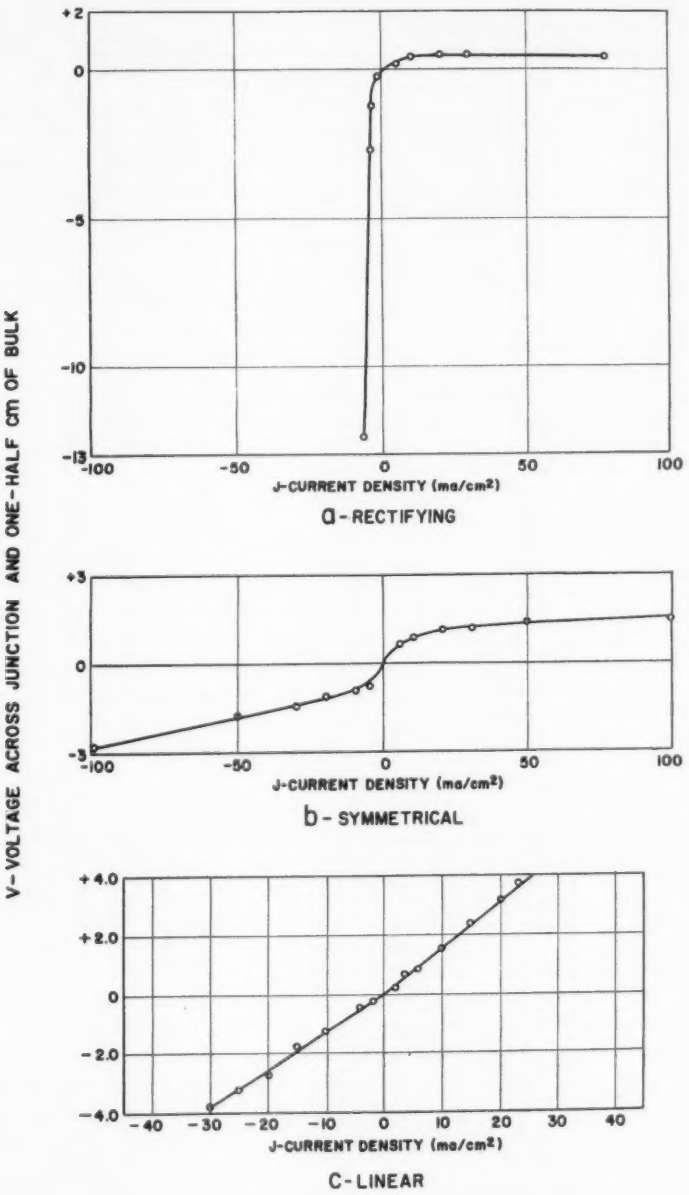


FIG. 6. Three current–voltage characteristics which may be observed for metal-to-semiconductor contacts. The symmetrical and linear characteristics occur for contacts which exhibit injection and extraction, respectively, for either direction of current flow. It is evident that the current–voltage characteristics cannot be used to make sensitive studies of such contacts.

surface. When aluminum was evaporated onto a cleaved surface which was exposed to the atmosphere, the surface was found to be still p-type but much less strongly so.

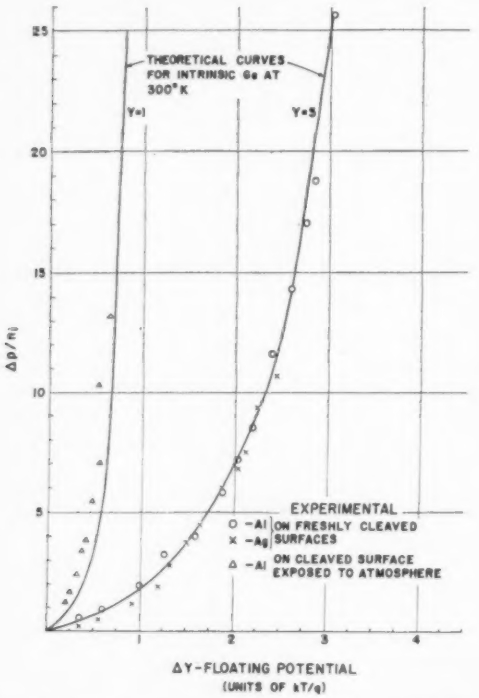


FIG. 7. Plots of contact floating potential versus added carrier density for a cleaved germanium surface. When the experimental results are compared to theoretical calculations, the barrier height under the contact can be determined. The surface is found to be p-type for both the fresh and the contaminated surfaces.

The techniques described in this section should be useful to workers studying electrolyte–semiconductor interfaces.

(2) *Absorption coefficients and spectrum of the carriers in the space charge region*

A knowledge of the absorption coefficients in the space charge region should yield some information about the scattering processes in the semiconductor surface and thus about the surface mobilities. The absorption coefficients can be measured through the field effect experiment where a known added charge is induced in the semiconductor surface. The measurements would, of course, have to be made quickly to avoid uncertainties arising from absorption of charge by the surface states.

The infrared spectrum of the holes is now explained in terms of the semiconductor band structure. It should also be possible to measure any change of this spectrum in going from the bulk to the surface and thus shed some light on the possible change in band structure.

(3) *Monitor of the carrier densities in the space charge region*

When the absorption coefficients are known, infrared radiation can be used to monitor surface conductivity, in the manner shown in Fig. 4c, as the carrier density in the space charge region changes. These changes may be brought about by a change in the surface states from chemical effects, by externally applied fields through the so-called field effect, by changing the gaseous ambient, or by bringing a metal with a different work function from the semiconductor in close proximity to the semiconductor surface. That such measurements can be made has already been demonstrated qualitatively (see Fig. 8), where it was shown, in particular, that the metal to semiconductor contact potential has an effect similar to that of an externally applied field in causing transient changes in the surface barrier height.[14] With a knowledge of the absorption coefficients and with the aid of a fast system, quantitative measurements can be made. Care needs to be taken in these measurements that changes in the surface reflectivity are properly considered.

(4) *Reflectivity measurements used to determine surface properties*

(a) *Internal reflectivity measurements.* Internal reflectivity measurements can be made using the same scheme as used to measure carrier densities in the semiconductor surface from absorption measurements as shown in Fig. 4(c). This technique has already been utilized to determine the reflectivity of a Ge–Hg interface as a function of angle of incidence at a wavelength of 3μ.[15] The results are shown in Fig. 9. It should be noted that a discrepancy exists between experiment and theory for the reflection coefficient of the parallel component of polarization, $R_{\parallel}$, at the principal angle. If

this discrepancy can be cleared up, measurements of $R_{\parallel}$ near its minimum may be used to give some information about the surface barrier. Because of the generally large reflectivity loss from metal–semiconductor interfaces, it is expected that little information can be gained about surface barriers from measurements of reflectivity before and after a metal comes into contact.

extreme caution must be taken with regard to signals from a change in absorption and recombination radiation. For the experimental set-up shown in Fig. 4(c), any changes in barrier height arising from changes in the gaseous ambient, field effect, or field from the contact potential could be qualitatively accounted for in terms of carrier absorption. Some observations attributed to changes in reflec-

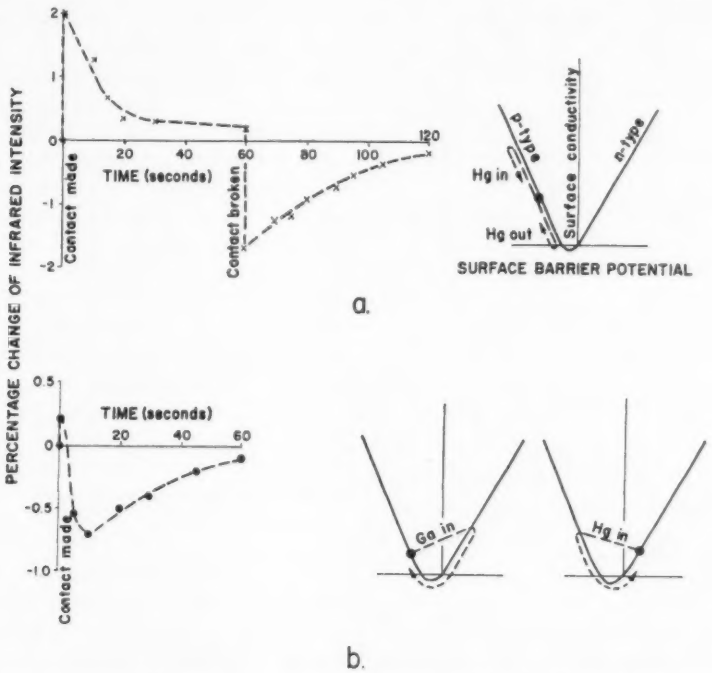


FIG. 8. Relaxation effects in the semiconductor surface arising from bringing a metal into contact with a semiconductor. These results indicate that the contact potential has an effect similar to that arising from an externally applied field as in the field effect experiment.

It is unlikely that internal reflectivity measurements from semiconductor–air surfaces can be used to study surface barriers because the modulation due to absorption is some 10^5 times greater, although it is of opposite sign. We have tried to make such measurements but did not observe any signals which could be attributed to reflectivity changes. Referring to Fig. 2, modulation by injection was found from beam I_{R2} but not for I_{R1}. The modulation in I_{R2} could be explained entirely by the change in the absorption for the internally reflected beam. For such measurements

tivity with carrier injection have been reported.[16] It is difficult to evaluate them, however, without more information.

(b) *External reflectivity measurements.* JOHNSON and MILLER[17] have attempted to investigate surface properties of semiconductors through the use of polarized light and making precise measurements of the external reflectivity coefficient for the parallel component of polarization near the principal angle. Their technique can be understood with reference to Fig. 9 which shows how the reflection coefficients for perpendicular and parallel

polarization change with angle of incidence for a germanium–mercury interface. Reflectivity theory indicates that the reflection coefficient of the parallel component of polarization at the principal angle is zero for an insulator. Any departure from

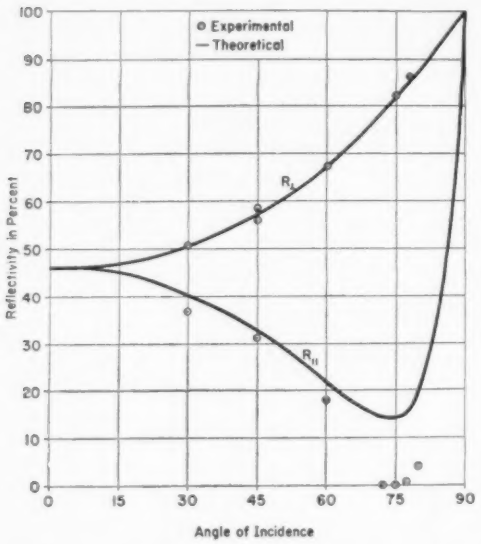


FIG. 9. The reflectivity coefficients for the perpendicular and parallel components of polarization for a germanium mercury interface as a function of angle of incidence at a wavelength of 3 μ.

zero should then be a measure of the surface conductivity. This approach thus seems to be a sensitive one because no large initial signal needs to be balanced off. In analysis of their measurements on rutile and strontium titanate they found it impossible to distinguish between a gradual change in carrier density, as may be found in a semiconductor surface, and an abrupt change, as may result from the presence of an oxide or other impurity layer on the surface. Thus it appears difficult to make absolute measurements in this way. However, their technique appears to have real merit if applied to semiconductor surfaces where the barrier height or carrier density in the barrier may be changed. The barrier height can be controlled through a field effect type of experiment. Carriers may be added to the surface by injection. It is here that the concentration of added carriers by magnetic deflection may really prove to be of value. Presumably large changes in carrier densities can be produced. From the calculated carrier distribution and a measure of the total number, using the technique shown in Fig. 4(c), the actual density at the surface can be determined. This added carrier density can be modulated easily which will add to the sensitivity of the experiment. Results in this experiment should provide a sensitive check of reflectivity theory. This technique, if successful, should provide a powerful tool to monitor the properties of a semiconductor surface. One particularly useful application of such a technique would be in a check of the Boltzmann relationships which relate the semiconductor surface and the neighboring bulk. Such a check has already been made for a p–n junction.[4] In order to use the Boltzmann relationship in solving the Poisson equation for the semiconductor surface in steady state, some assumption has to be made; e.g. constancy of charge in the surface barrier. It is not clear to us that this assumption is completely justified. One could check such an assumption by determining the added carrier in the bulk, the change in carrier densities at the surface and simultaneously determining the surface photovoltage. Such an experiment is an important one as the relationship in question has bearing on metal to semiconductor contact characteristics and free surface theory.

CONCLUSION

Infrared radiation might be used to study semiconductor surfaces in one of the following ways: From a monitor of the added carrier density in the bulk or in the space charge region; through a study of the spectral characteristics of the carriers in the space charge region; or through a study of the reflectivity as a function of carrier density. Some of these applications have already been realized and have yielded useful information especially in the field of metal to semiconductor contacts where the nature of the surface is a controlling factor. It is shown that more sensitive tests of the contacts can be made by examining the current versus added carrier density characteristics rather than the more usual current–voltage characteristics. It is shown that recombination velocity and barrier height under the contact can be measured. These and other experiments open up a new line of attack to determine the nature of the semiconductor surface.

VOL.
14
1960

REFERENCES

1. FAN H. Y. and BECKER M. studied the absorption characteristics of equilibrium carriers in germanium, see *Semiconducting Materials*, p. 132, Butterworths, London (1951). The mechanism of absorption of infrared radiation by the conduction electrons in *n*-type germanium is discussed by MEYER H. J. G., *Phys. Rev.* **112,** 298 (1958). GIBSON A. F. has discussed the application of Drude–Zener theory to absorption of infrared radiation by the free carriers in semiconductors, *J. Sci. Instrum.* **35,** 273 (1958).
2. The use of semiconductors with junctions as infrared modulators was first described by LEHOVEC K., *Proc. Inst. Radio Engrs., N.Y.* **40,** 1407 (1952). The spectral absorption characteristics of the added free carriers was investigated almost simultaneously by BRIGGS H. B. and FLETCHER R. C., *Phys. Rev.* **91,** 1342 (1953); NEWMAN R., *Phys. Rev.* **91,** 1311 (1953); GIBSON A. F., *Proc. Phys. Soc., Lond.* B**66,** 588 (1953), and more recently by KESSLEN F. R., *J. Phys. Chem. Solids* **8,** 275 (1959).
3. HARRICK N. J., *Phys. Rev.* **103,** 1173 (1956); HULDT L. and STAFLIN T., *Phys. Rev. Letters* **1,** 236 (1958) and UKHANOV Y. I., *Zh. tekh. fiz.* **28,** 2410 (1958) [English Translation **3,** 2212 (1959)] have been using this technique to probe out carrier distributions.
4. HARRICK N. J., *J. Appl. Phys.* **29,** 764 (1958).
5. HARRICK N. J., *J. Appl. Phys.* **27,** 1439 (1956); HULDT L., *Phys. Rev. Letters* **2,** 3 (1959).
6. HARRICK N. J., to be published.
7. BEYEN W., BRATT P., DAVIS H., JOHNSON L., LEVINSTEIN H. and MACRAE H., *J. Opt. Soc. Amer,* **49,** 686 (1959).
8. SHOCKLEY W., *Electrons and Holes in Semiconductors*, p. 325. Van Nostrand, Princeton, N.J. (1950).
9. WEISSHAAR E. and WELKER H., *Z. Naturf.*, **8a,** 681 (1953); MADELUNG O., TEWORDT L. and WELKAR H., *Z. Naturf.* **10a,** 476 (1955); PIKUS G. E. and SOROKIN O. V., *Soviet Phys. Technical Physics*, 2457 (1957).
10. HARRICK N. J., to be published.
11. HARRICK N. J., *Phys. Rev.* **115,** 876 (1959).
12. HARRICK N. J., *Phys. Rev. Letters* **2,** 199 (1959).
13. JOHNSON E. O., *Phys. Rev.* **111,** 153 (1958).
14. HARRICK N. J., *J. Phys. Chem. Solids* **8,** 106 (1959).
15. HARRICK N. J., *J. Opt. Soc. Amer.* **49,** 376 (1959).
16. FILINSKI I., *Phys. Rev.* **107,** 1193 (1957).
17. MILLER P. H., JR. and JOHNSON J. R., *Physica,'sGrav.* **20,** 1026 (1954).

JOHNSON J. R., unpublished Ph.D thesis, University of Pennsylvania (1958); Tech. Rep. No. 15 for Contract N6-onr-24914.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 72–74. Printed in Great Britain.

MEASUREMENTS OF SURFACE TRANSPORT PHENOMENA

T. H. GEBALLE

Bell Telephone Laboratories, Murray Hill, New Jersey

TRANSPORT properties in bulk germanium have been carefully studied and are quite well understood in many cases. The use of strain fields and thermal gradients as well as electric and magnetic fields over a wide range of temperature has been valuable. In the present paper the extension of some of these variables to measurements of surface phenomena will be considered.

In particular, the thermoelectric power, Q, would appear to be a promising quantity to investigate. The experimental measurement should require only minor modification of the existing techniques that are used in measuring surface conductance. The only additional variable to be determined is the thermal gradient which can be estimated from the measured Seebeck voltage and the known Q for bulk germanium in a manner similar to that which has been done in investigations of the thermomagnetic properties of bulk germanium.[1] Thus it is not even necessary to add thermocouples to the sample. The measured Q is simply the average of Q_s and Q_b weighted with respect to their conductances g_s and g_b, where the subscripts refer to surface and bulk. If Q_s and Q_b are not equal then circulating currents will flow and the measured Q will be given by the following.

$$Q_{\text{meas}} = (Q_s g_s + Q_b g_b)/(g_s + g_b). \tag{1}$$

If one applies a variable electric field ϵ at the surface, Q will be modulated according to the following relation.

$$\left(\frac{dQ}{d\epsilon}\right)_{\text{meas}} \sim \frac{1}{g_b}\left(Q_s\frac{dg_s}{d\epsilon} + g_s\frac{dQ_s}{d\epsilon}\right) \tag{2}$$

assuming $g_s \ll g_b$. Working around the region of surface potential where $dg_s/d\epsilon$ is large, the first term in the brackets should dominate so that (2) should reduce to a simple expression for Q_s in terms of measureable quantities. However, the interpretation of Q_s in terms of a microscopic model will require the same order of theoretical work as does the calculation of mobility and other transport parameters. Above room temperature Q_s depends upon the location of the Fermi level and a scattering term.

$$Q_s = (\Delta\epsilon_F)_s + AkT. \tag{3}$$

A depends on the type of scattering and distribution of carriers and is of order unity; and $(\Delta\epsilon_F)_s$ is the energy difference between the Fermi level and the appropriate band edge properly averaged over direction normal to the surface.

Below room temperature phonon-drag effects may become important. Q is enhanced in value above that given by the tendency of the mobile carriers to diffuse in the thermal gradient due to the preferential scattering of the carriers by the asymmetric distribution of phonons which exists of necessity in the thermal gradient. The phonon-drag contribution to Q is proportional to the strength of the phonon–electron coupling and to the relaxation time of the long wavelength phonons with which the electrons interact. Thus the phonon drag contribution Q_s will be sensitive to the degree of boundary scattering of the phonons. Such sensitivity of Q, and indeed of the thermal conductivity also, has been observed at low temperatures where boundary scattering is the chief mechanism for returning the phonons to thermal equilibrium. It has been found that at low temperatures etched (111) planes exhibit a higher degree of specular reflection of phonons than other etched surfaces, thus indicating that such planes look smoother to phonons of a few hundred angstroms wavelength than (110) and (100) planes.

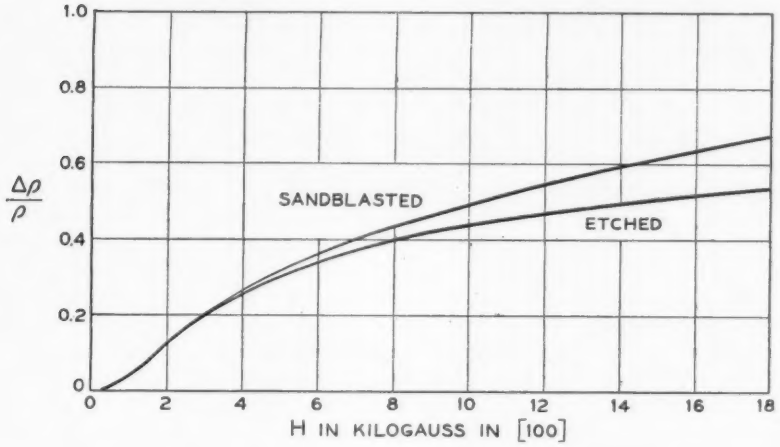


FIG. 1. Transverse magnetoresistance in high purity *n*-type germanium at 77°K for two different sample preparations. The sample cross-section is approximately 1 mm × 4½ mm. (After HERRING, GEBALLE and KUNZLER[1].)

Perhaps the surface mobility of electrons and holes may also depend on the plane of the surface on which the measurements are taken.

It has been possible to obtain a wealth of detailed quantitative information in bulk germanium by superimposing a magnetic field on the thermal gradient and studying thermomagnetic phenomena.[1] At the stage of development of surface physics at the present time where the simpler quantities such as conductivity itself have not been thoroughly investigated over a wide range of temperature it is perhaps premature to become too

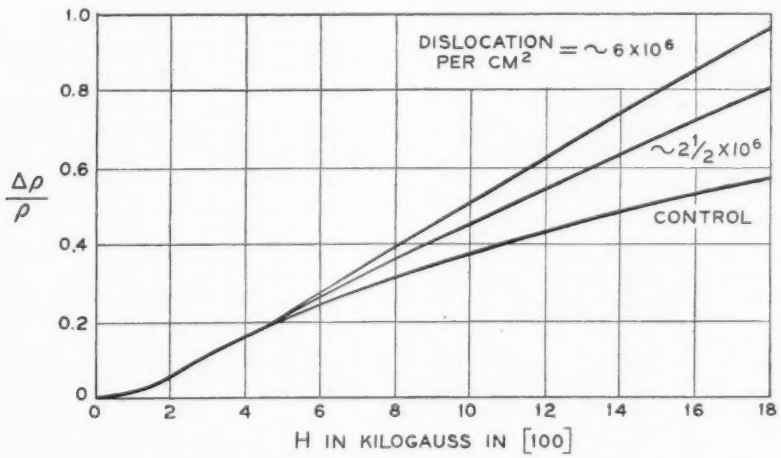


FIG. 2. Transverse magnetoresistance in high purity *n*-type germanium at 77°K for samples which were heated and plastically deformed (through the courtesy of R. A. Logan) so as to produce the concentrations of (1TO) dislocations indicated. The control sample was heated but not purposely bent. (After GEBALLE T. H. and KUNZLER J. E., unpublished.)

involved in speculating how the more complicated transport properties behave.

However, rough observations already made indicate a sensitivity to surface conduction of quantities such as the Nernst effect where in some temperature regions the bulk effect becomes zero so that only surface effects are observed. The Nernst field, which is the analogue of the Hall field when a thermal gradient replaces the applied electric field, is the resultant of the competition between the Hall fields of different groups of carriers which diffuse differently in the thermal gradient. In high-purity *n*-type germanium diffusion of the different groups of electrons becomes either independent of energy or at least a non-monotonic function of energy around 200°K due to the counteracting influences of the phonon drag and the normal kinetic terms. As a result, the bulk effect becomes very small and the influence of the surface becomes important as is evident from the observation that the sign of the Nernst field reverses upon sandblasting a previously etched surface. The presence of light and heavy mass holes in *p*-germanium and *p*-indium antimonide is strikingly evident in measurements of the Nernst effect as a function of magnetic field in the phonon-drag region,[2] even though a quantitative analysis is not easy. The presence of light and heavy holes in surface inversion regions might also be easy to detect.

The change of surface conduction with applied strain, or the surface piezoresistance, is another effect which presumably can be measured. While such measurements have been valuable in determining the band structure of bulk germanium, it is not obvious that they could be made accurately enough to be of value in determining surface properties.

Other effects such as the high-field transverse magnetoresistance may be of more value. In high-purity germanium at liquid nitrogen temperatures, fields of 3000 G cause Hall angles of the order of one radian to exist in the bulk material. The large Hall fields which occur in higher magnetic fields can be shorted by surface conduction by dislocations or by less well-defined inhomogeneities. Surface conductance of the order of one per cent of the bulk conductance is presumably the cause of serious nonsaturation of the observed transverse high-field magnetoresistance in at least some cases. Fig. 1 shows typical data for a high-purity sample whose bulk properties were carefully investigated. The gross effect of eliminating the conduction paths in a sandblasted surface is clearly evident. In Fig. 2 a similar effect can be seen in germanium with etched surfaces, but plastically deformed so as to produce an array of dislocations parallel to the Hall field. If techniques are developed so that high magnetic fields can be applied to clean surfaces, transverse magnetoresistance measurements should be useful in studying surface transport.

REFERENCES

1. Herring C., Geballe T. H. and Kunzler J. E., *Phys. Rev.* **111,** 36 (1958); *Bell Syst. Tech. J.* **38,** 657 (1959).
2. Herring C., *Halbleiter und Phosphore* (edited by Schön M. and Welker H.) p. 228. Vieweg, Braunschweig (1958).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 75–76. Printed in Great Britain.

STRUCTURES OF MONOLAYERS OF ADSORBED GASES

L. H. GERMER and C. D. HARTMAN

Bell Telephone Laboratories, Murray Hill, New Jersey

VOL. 14 1960

DIFFRACTION patterns of low energy electrons produced by a crystal surface have been accelerated sufficiently to show on a fluorescent screen where they can be photographed.[1] Various gases at low controlled pressures have been admitted to a clean surface, and the arrangements of atoms or molecules upon the surface deduced from their diffraction patterns.[2] In the first application of the method a (111) face of a nickel crystal was used. the order of 1000 volts to a fluorescent screen. The pattern is observed on the screen, or photographed by its reflection in a mirror inclined by 45° to the primary beam direction.

The tube of Fig. 1 is attached to vacuum pumps, and a pressure of 2×10^{-10} mm Hg of residual gas can be attained. Pure gases at low pressures can be admitted through a bank of six Granville–Phillips type vacuum valves. The gases studied up to this

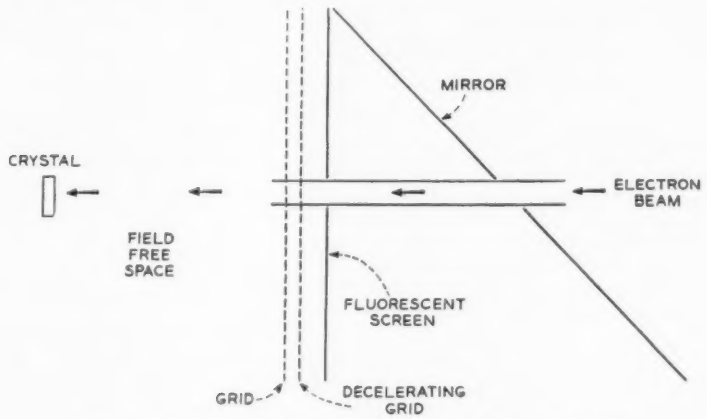


FIG. 1. Schematic drawing of the electron diffraction tube.

It is proposed to apply it to gases adsorbed upon semiconductor surfaces, and for this reason presentation at this Conference seems in order. However, the work reported here is concerned only with gases adsorbed upon a (111) face of a nickel crystal.

The experimental tube is shown schematically in the sketch of Fig. 1. Electrons emerging from the gun strike the crystal at normal incidence. The diffracted electrons move in a field free space until they reach a fine-mesh grid. Between this grid and a second grid is a decelerating field to remove most of the slow electrons. Beyond the second grid the diffracted electrons are accelerated by a voltage of time have been nitrogen, carbon monoxide, oxygen and hydrogen.

Molecules of CO, which behave most simply of the four gases, are adsorbed on a (111) nickel face as a monolayer with the spacing between adjacent molecules double that between neighboring nickel atoms in the surface. The arrangement is that marked (4) in Fig. 2, this number representing the ratio of the number of surface nickel atoms to the number of CO molecules. The sticking probability is high, unity according to the crude method by which we have measured it. The sharpness of the diffraction patterns is that

expected from a single two-dimensional crystal. The CO is completely removed by heating the crystal to 175°C.

The behavior of nitrogen is much more complex. At low pressures it is adsorbed, with a quite low sticking probability and apparently as atoms, in a

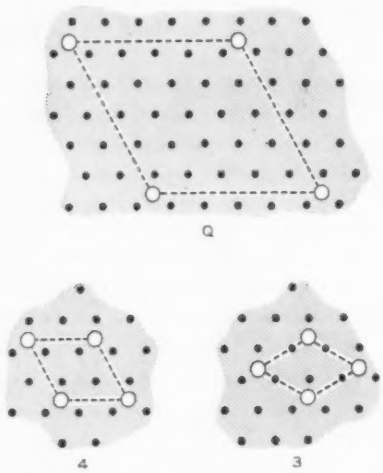


FIG. 2. Sketches of the three different arrangements which atoms and molecules have been found to take up upon a (111) face of a nickel crystal. Black dots are surface nickel atoms, and open circles adsorbed atoms or molecules. The numbers (3) and (4) represent the ratio of the numbers of surface nickel atoms to the numbers of adsorbed atoms or molecules. Location of atoms within the unit cell of the Q-structure is variable.

monolayer having this same two dimensional crystalline structure. Diffuseness of the diffraction spots indicates, however, that the layer is polycrystalline. Crystal growth can be induced by heating to 125°C. Heating about 200°C results in transformation into one or the other of two more stable polymorphic forms, the unit cells of which are marked 3 and Q in Fig. 2. The 3-structure is completely identified in the figure, but the locations of the nitrogen atoms within the cell of the Q-structure are different in different cells and the distribution of positions is different in different experiments. Heating above 300°C results in melting the nitrogen monolayer without removing it, the crystalline structure returning upon cooling. The nitrogen is completely removed at temperatures somewhat above 400°C.

Oxygen can be adsorbed in a layer several atoms or molecules thick. This shows no structure until annealed at 125°C after which very diffuse diffraction spots appear representing arrangement in the Q-structure. The diffraction spots can be considerably sharpened by further anneal. What is apparently a single monolayer of oxygen gives the same diffraction pattern, after it, too, has been annealed. (This pattern is not that of normal NiO.)

No diffraction patterns have been obtained from hydrogen nor does hydrogen produce any weakening of the normal Laue beams from the underlying nickel crystal. The presence of hydrogen on the crystal surface does, however, alter the nickel diffraction beams in a peculiar way; the voltage discrimination of the normal Laue beams is much improved by the presence of hydrogen on the crystal surface, as if the resolving power of the crystal were improved by surface hydrogen. This effect is shown by oxygen also.

Among proposed continuations of the present experiments, we expect to study gases adsorbed upon a smooth single crystal silicon surface, and to investigate catalytic reactions between different gases upon surfaces of nickel and silicon.

REFERENCES

1. This technique was first used by EHRENBERG W., *Phil. Mag.* **18,** 878 (1934), who, however, did not develop it sufficiently to make it a practically useful research tool.
2. Surface arrangements of adsorbed gas atoms or molecules were first determined by electron diffraction in 1927—DAVISSON C. J. and GERMER L. H., *Phys. Rev.* **30,** 705 (1927); GERMER L. H., *Z. Phys.* **54,** 408 (1929). The method has since been extensively developed by FARNSWORTH H. E. —many papers from *Nature, Lond.* **123,** 941 (1929) to *J. Appl. Phys.* **30,** 675 (1959). Up to the present time all work in this field has been by slowly exploring the diffraction pattern by means of a Faraday box collector. With the new photographic technique, experimentation has been made very much simpler and experiments can be carried out which were heretofore impossible.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 77–86. Printed in Great Britain.

SECTION III. CHEMISTRY

SOME ASPECTS OF THE CHEMISTRY OF GERMANIUM AND SILICON SURFACES

MINO GREEN

Zenith Radio Corporation, Chicago 39, Illinois

VOL. 14 1960

Abstract—Some recent aspects of the chemistry of germanium and silicon surfaces are reviewed and discussed. The paper is divided into three parts: clean surfaces, adsorption from the gaseous phase and electrochemistry.

1. INTRODUCTION

THIS review of some recent aspects of the chemistry of germanium and silicon surfaces is organized under three main headings: clean surfaces, adsorption from the gaseous phase, and electrochemistry. Germanium is dealt with more fully than silicon, a state of affairs which represents our relative state of knowledge of these substances.

2. CLEAN SURFACES

An understanding of the chemical properties of clean semiconductor surfaces requires, among other things, a knowledge of both the electronic properties of the surface, and the dominant geometrical features of the surface, such as the main crystallographic arrangement of the surface atoms and what kind of important non-ideal features are present, e.g. micro-cleavage steps.

During the last few years it has been clearly established[1–3] that a clean germanium surface, prepared by the argon-ion bombardment/annealing cycle, has a highly electrically conducting *p*-type skin. It has been assumed, reasonably, that the *p*-type skin arises because of the presence of a high excess of bound surface electrons. The excess of surface electrons is brought into being by surface states.[4] It has not, however, yet been established beyond any reasonable doubt that these surface states are a property of a hypothetically dislocation-free surface. The alternative is that surface states on germanium might be intimately associated with step edges, kink sites and vacant sites on the crystal surface. With this in mind, it will be interesting to measure the surface conductivity of germanium surfaces prepared in a different way from the argon-ion bombardment method; hopefully the extent of surface imperfection would be different. An alternative cleaning method is available.[5] It has been shown that, if germanium is etched so as not to give a thick oxide layer and the germanium is then heated in a high vacuum to ~700°C, a clean germanium surface results. There are several etching methods that can be used; one method is simply to etch with CP-4 and then to continuously dilute the etch with water, on no account allowing the germanium to be exposed to air until the etch has been thoroughly washed away.

Studies of the crystallography of clean germanium[6, 7] and silicon[8] surfaces have shown that the surface atoms in crystal planes of low index are not in their ideal positions. The deviation from ideality takes the form of an ordered pairing of adjacent surface atoms in such a way as to give rise to a surface grating with double spacing, as well as the expected single spacings. In addition to the slight atomic displacements just mentioned, it would be of great interest to know more about the nature and extent of gross imperfections on germanium and silicon surfaces. For example, WOLFF and BRODER[9] have shown by electron-micrography that a fractured silicon surface is far from atomically flat.

3. ADSORPTION FROM THE GASEOUS PHASE

A theory of adsorption on semiconductor surfaces is only slowly coming into being, and the form that it seems to be taking is best indicated by the work of KOUTECKY[10] and GRIMLEY[11]. A recent tendency to an *ad hoc* application of MULLIKEN's theory[15] of "molecular compounds" to adsorption on semiconductor surfaces[16] seems destined to fail as a general theory under a more searching examination. The "boundary layer theory of adsorption"[12–14] is not applicable to covalent type semiconductors.

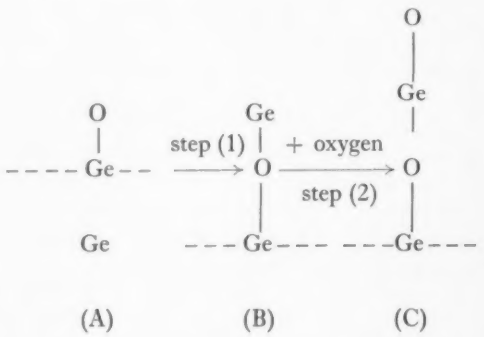
Information on the adsorptive properties of clean germanium and silicon surfaces continues to appear. The now almost "classical" germanium–oxygen system will be discussed in some detail and briefer mention given to the other less well-studied systems.

The general features of the adsorption of O_2 on clean germanium surfaces in the temperature range $-78°C$ to $\sim 80°C$ are as follows:[17, 18] There is an initially very rapid adsorption of O_2 (sticking factor[19] $\sim 0{\cdot}002$) up to approximately two-thirds monolayer coverage, and a subsequent slower adsorption associated with the formation of a hyperthin oxide layer. The kinetics of the slower adsorption follows the logarithmic rate expression

$$N = a + b \log_{10} t,$$

where N is the amount taken up at time t, and a and b are temperature dependent constants; a is also pressure dependent. For about the same temperature and pressure, various workers[17, 20] have reported values of the ratio a/b ranging from 5 to 25. DELL[20] was the first to point out that a and b could be structure sensitive. Preliminary studies by GREEN and LIBERMAN[21] using the drop hammer method support the idea that a and b are structure sensitive constants and would suggest that the ratio a/b depends on the initial mechanical condition of the crystal as well as the extent to which a crystal is crushed. A spurious, but experimentally serious, source of variation in a and b is partial diffusion control of the adsorption process, i.e. some portion of the powder is for one reason or another only accessible through very fine pores. Partial diffusion control of oxygen take-up on germanium crushed by the drop hammer method has been observed in some cases.[5] It is apparent that, when adsorption kinetics are measured on systems where the adsorbent is a finely divided powder, great care should be exercised to eliminate, where possible, rates due to slow gas diffusion. The easiest way to find out if gaseous diffusion is important in germanium or silicon systems is to agitate the powder and observe the effects, if any, on the overall adsorption kinetics.

The nature of adsorbed oxygen below monolayer coverage, and before the onset of oxide formation, is not clear. The heat of adsorption[17, 21] is such as to suggest adsorption in the form of atoms,[18, 22] and if this is so, oxygen on certain germanium surfaces must still be a free radical, i.e. ($\geq$Ge–O). Peroxide formation has been suggested,[2] i.e. Ge–O–O–Ge; however, this seems unlikely when the heat of adsorption and steric effects are considered together. The experiments of HANDLER *et al.*[2, 3] and others,[1] show that initially the adsorption of oxygen causes the surface to go more p-type. The maximum in the p-type surface conductivity would appear to occur at about one-half to one monolayer of oxygen atoms.[23] This is consistent with the idea that, for this range of coverage, oxygen adsorbs as covalently bound free radicals. HANDLER and PORTNOY[24] have reported that, when the oxygen in their adsorption chamber has been pumped off, there is a partial reversal in the curve of the surface conductance as a function of oxygen exposure. A possible explanation of this observation can be based on the mechanism of oxide growth. Germanium, covered with oxygen atoms, is converted (in the presence of oxygen gas) to germanium covered with a hyperthin oxide by some sort of surface rearrangement,[18] e.g.



Step (1) can occur in the absence of oxygen (this has been observed for the Si/O_2 system[25]). In the

VOL. 14 1960

presence of oxygen there will be a certain proportion of surface species (A), (B) and (C) depending upon the extent of oxidation. The surface conductance will depend upon the concentration of these surface species. If, however, oxygen is pumped off, step (1) will still go on and, for a given oxygen take-up the concentration of surface species (B) will increase. This may account for a partial reversal of surface conductance versus oxygen exposure curve without having to invoke the necessity of oxygen desorption.

Recent studies of the adsorption of oxygen on silicon are not in accord. Thus LAW[25], working on clean evaporated films of silicon, reported kinetics similar to the Ge/O_2 kinetics. On the other hand, GREEN and MAXWELL[26] found, for crushed powders, that there was an initial fast adsorption followed by a slower adsorption which was not logarithmic, but which came to a complete halt at a take-up of about 1·5 oxygen atoms/surface silicon atom. Clearly, further experimental work is required for the Si/O_2 system.

Studies of the interaction of hydrogen with clean germanium surfaces show that: molecular hydrogen does not absorb[27]; atomic hydrogen (made by high-temperature thermal decomposition) adsorbs to the extent of about a monolayer[3, 6]; on evaporated films of germanium para-hydrogen is readily converted to ortho-hydrogen[28] and the H_2–D_2 exchange occurs.[28] Surface conductivity experiments suggest that adsorbed hydrogen behaves in an unusual way. Thus, the adsorption of atomic hydrogen causes an increase in the *p*-type surface conductivity.[3, 24] But the elimination of unpaired surface electrons by the formation of simple covalent bonds, viz. $\rightarrow$Ge–H or $\rightarrow$Ge$\begin{smallmatrix}\diagup\\\diagdown\end{smallmatrix}\begin{smallmatrix}H\\H,\end{smallmatrix}$ is, somewhat crudely, associated with the elimination of surface states, and the loss of enhanced surface conductivity. When germanium, covered with a monolayer of hydrogen, is exposed to oxygen, the oxygen is adsorbed but hydrogen is not displaced[29] nor is water formed. This suggests that adsorbed hydrogen atoms may be "buried" in the germanium surface and not held to the surface by a simple covalent bond. Atomic hydrogen adsorbs to the extent of a monolayer on clean silicon surfaces,[30] while molecular hydrogen adsorbs to a negligible extent. EISINGER[31] has measured the effect of adsorbed atomic hydrogen on the photoemission from silicon surfaces. An increase in photothreshold of 0·4 eV was observed for monolayer coverage.

The adsorption of a number of readily condensable gases on clean germanium (crushed crystal) surfaces has recently been studied.[27] At room temperature and for a relative pressure of $\lesssim 5\times10^{-3}$, the gases found to adsorb were H_2O, CH_3OH, $(CH_3)_2CO$, CH_3COOH, $(CH_3)_3CNH_2$, C_5H_5N, and *p*-dioxan, while C_2H_4, CCl_4, $CHCl_3$, $(C_2H_5)O_2$, C_6H_6 and C_6H_5Cl were found not to adsorb. For those gases that were adsorbed, the adsorption was extremely rapid and the extent of take-up was equivalent to about monolayer coverage. The conclusion drawn from this study was that the adsorption or non-adsorption of all the gases mentioned above, excepting *p*-dioxane and pyridine, could be explained in terms of the electrostatic interaction between the permanent charge distributions of the gas and the germanium surface. The nature of the binding, for these systems where adsorption occurred, was described as a form of "hydrogen bond".

A study of the adsorption of the inert gases on germanium and silicon is of interest because of the information it can give about the nature of the adsorbent. Two systems have been examined in detail. These are: Ge–Kr studied by ROSENBERG[32] and Ge–A studied by SPARNAAY[33]. Rosenberg concluded from his work that "the energy of krypton adsorption cannot be accounted for in terms of the atomic properties of bulk germanium". A reappraisal of ROSENBERG's data by GREEN and SEIWATZ[34] does not substantiate this conclusion, indeed these authors conclude that the measured heat of adsorption of krypton on germanium is consistent with the properties of crystalline germanium and krypton atoms. It is interesting to note, in connection with this latter finding, that PIEROTTI and HALSEY[35] have examined the adsorption of krypton on a number of metals and conclude that the adsorption forces are readily explained in terms of the atomic properties of the metal.

4. ELECTROCHEMISTRY

Studies of the electrochemistry of the germanium–electrolyte electrode (silicon has not been studied enough to warrant comment) can be roughly divided into two groups, depending upon

whether the electrode was contaminated or not. The problem of surface contamination in electrochemical cells[36] is every bit as severe as the "clean surface" problem in gas adsorption.

BRATTAIN and GARRETT[37] published the first comprehensive study of the contaminated germanium electrode. The virtue of this work was that it clearly demonstrated that electrolysis could, under suitable conditions, be governed by the rate of diffusion of minority carriers to the germanium–electrolyte interface. Since the surfaces were contaminated, nothing could be deduced about the mechanism of discharge. Several further studies of the germanium–electrolyte electrode have appeared[38–40] which confirm and extend BRATTAIN and GARRETT's observations.

The hydrogen evolution reaction at germanium surfaces in clean electrochemical systems has been studied by a number of workers.[41–45] It has been concluded[43, 45] that the proton discharge mechanism, when semiconductor minority carrier diffusion control is not rate determining, is

$$H_3O^+ + e^- \xrightarrow{\text{fast}} H(\text{adsorbed}) + H_2O$$

$$H_3O^+ + H(\text{adsorbed}) + e^- \xrightarrow{\text{slow}} H_2(\text{gas}) + H_2O.$$

It was observed that the kinetics of the hydrogen evolution reaction were independent (below the range of semiconductor-free carrier diffusion control) of the concentration and type of free carriers in the bulk semiconductor. However, the discharge kinetics were dependent upon the particular crystal face of germanium which was exposed to the electrolyte, cf. Table 1. The latter observations are to be expected from the theory of the semiconductor–electrolyte electrode.[46, 47]

Table 1. *Hydrogen overpotential data at germanium electrodes* (28°C and 1M *hydrochloric acid*)

Crystal surface exposed	Bulk electron concentration (cm^{-3})	Exchange current (A. cm^{-2})	$-b$(mV)*
unknown	4×10^{14}	$1{\cdot}4\times10^{-8}$	113
unknown	3×10^{12}	$4{\cdot}7\times10^{-6}$	102
(100)	$\sim10^{15}$	$(2{\cdot}2\pm0{\cdot}3)\times10^{-8}$	115 ± 12
(100)	$\sim10^{12}$	$(2{\cdot}3\pm0{\cdot}8)\times10^{-8}$	109 ± 5
(110)	$\sim10^{12}$	$(1{\cdot}0\pm0{\cdot}2)\times10^{-9}$	116 ± 2
(111)	$\sim10^{12}$	$(1{\cdot}9\pm0{\cdot}4)\times10^{-8}$	148 ± 9

* $b = 2{\cdot}303\ (d\eta/d\ln i)$ per decade of current, where η is the overpotential and i is the current density.

The behavior of the germanium–electrolyte electrode, with no net current passing, has been the subject of several investigations.[22, 48, 49] The dissolution of germanium has been observed, and was attributed to a "mixed-potential" type of reaction, viz.

(I) $Ge + 2H_2O \rightarrow GeO_2\,(\text{dissolved}) + 4H^+ + 4e^-$

(II) $4H^+ + 4e^- \rightarrow 2H_2(\text{gas})$

(III) $Ge + 2H_2O \rightarrow GeO_2\,(\text{dissolved}) + 2H_2\,(\text{gas})$

HARVEY[50] has shown that the overall dissolution reaction (III) goes extremely slowly, and that the germanium–electrolyte electrode behaves as a reversible hydrogen electrode. Deviations from hydrogen electrode like behavior can be attributed to trace quantities of oxidizing agent.

REFERENCES

1. AUTLER S. H., MCWHORTER A. L. and GEBBIE H. A., *Bull. Amer. Phys. Soc.* **1,** 145 (1956).
2. HANDLER P., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 23. University of Pennsylvania Press, Philadelphia (1957).
3. HEILAND G. and HANDLER P., *J. Appl. Phys.* **30,** 446 (1959).
4. BARDEEN J., *Phys. Rev.* **71,** 717 (1947).
5. GREEN M. and LIBERMAN A., unpublished work.
6. SCHLIER R. E. and FARNSWORTH H. E., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 3. University of Pennsylvania Press, Philadelphia (1957).
7. SCHLIER R. E. and FARNSWORTH H. E., *J. Chem. Phys.* **30,** 917 (1959).
8. FARNSWORTH H. E., SCHLIER R. E. and DILLON J. A., *J. Phys. Chem. Solids* **8,** 116 (1959).
9. WOLFF G. A. and BRODER J. D., *Acta Cryst.* **12,** 313 (1959).
10. KOUTECKY J., *Trans. Faraday Soc.* **54,** 1038 (1958).
11. GRIMLEY T. B., *Proc. Phys. Soc.* **72,** 103 (1958).
12. WEISZ P. B., *J. Chem. Phys.* **20,** 1483 (1952); **21,** 1531 (1953).
13. AIGRAIN P. and DUGAS C., *Z. Elektrochem.* **56,** 363 (1952).
14. HAUFFE K. and ENGELL H. J., *Z. Elektrochem.* **56,** 366 (1952).
15. MULLIKEN R. S., *J. Amer. Chem. Soc.* **64,** 811 (1952).
16. MATSEN F. A., MAKRIDES A. C. and HAKERMAN N., *J. Chem. Phys.* **22,** 1800 (1954).
17. GREEN M., KAFALAS J. A. and ROBINSON P. H., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 349. University of Pennsylvania Press, Philadelphia (1957).
18. GREEN M., *ibid.* p. 362.

19. DILLON J. A. and FARNSWORTH H. E., *J. Appl. Phys.* **28,** 174 (1957).
20. DELL R. M., *J. Phys. Chem.* **61,** 1584 (1957).
21. BRENNAN D., private communication.
22. GREEN M., *Progress in Semiconductors*, Vol. IV. p. 35. Heywood, London (1960).
23. SEIWATZ R., private communication.
24. HANDLER P. and PORTNOY W. H., *Phys. Rev.* **116,** 516 (1959).
25. LAW J. T., *J. Phys. Chem. Solids* **4,** 91 (1958).
26. GREEN M. and MAXWELL K. H., *J. Phys. Chem. Solids*, to be published.
27. GREEN M. and MAXWELL K. H., *J. Phys. Chem. Solids* **11,** 195 (1959).
28. SANDLER Y. L. and GAZITH M., *J. Phys. Chem.* **63,** 1095 (1959).
29. MAXWELL K. H. and GREEN M., *J. Phys. Chem. Solids*, to be published.
30. LAW J. T., *J. Chem. Phys.* **30,** 1568 (1959).
31. EISINGER J., *J. Chem. Phys.* **30,** 927 (1959).
32. ROSENBERG A. J., *J. Phys. Chem.* **62,** 1112 (1958).
33. SPARNAAY M. J., *Solid State Physics in Electronics and Telecommunications*, Vol. I, pt. I. Academic Press, London (1960).
34. GREEN M. and SEIWATZ R., unpublished work.
35. PIEROTTI R. A. and HALSEY G. D., *J. Phys. Chem.* **63,** 680 (1959).
36. BOCKRIS J. O'M. and CONWAY B. E., *Trans. Faraday Soc.* **45,** 989 (1949).
37. BRATTAIN W. H. and GARRETT C. G. B., *Bell. Syst. Tech. J.* **34,** 129 (1955).
38. EFIMOV E. A. and ERUSALIMCHIK I. G., *Zh. fiz. khim.* **32,** 413 (1958).
39. TURNER D. R., *J. Electrochem. Soc.* **103,** 252 (1956).
40. FLYNN J. R., *J. Electrochem Soc.* **105,** 715 (1958).
41. GREEN M., *Lincoln Laboratory Group Report* M35–60 (Oct. 1956).
42. BOCKRIS J. O'M., private communication.
43. GREEN M., *Solid State Physics in Electronics and Telecommunications*, Vol. I, pt. 1, p. 619. Academic Press, London (1960).
44. EFIMOV E. A. and ERUSALIMCHIK I. G., *Zh. fiz. khim.* **32,** 1967 (1958).
45. ERDELYI V. R. and GREEN M., *Nature, Lond.* **182,** 1592 (1958).
46. GREEN M., *Modern Aspects of Electrochemistry*, Vol. II. Butterworths, London (1959).
47. GREEN M., *J. Chem. Phys.* **31,** 200 (1959).
48. HARVEY W. W. and GATOS H. C., *Trans. Electrochem. Soc.* **105,** 654 (1958).
49. LOVRECEK B. and BOCKRIS J. O'M., *J. Phys. Chem.* **63,** 1368 (1959).
50. HARVEY W. W., private communication.

VOL. 14 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 82–86. Printed in Great Britain.

CHEMICAL CONFIGURATION OF GERMANIUM SURFACES IN AQUEOUS MEDIA

W. W. HARVEY

Lincoln Laboratory,* Massachusetts Institute of Technology, Lexington 73, Massachusetts

Abstract—An essentially electrochemical technique has been employed for measuring changes in surface recombination velocity v_s for an n-type germanium electrode. Displacement of dissolved oxygen from the solution leads to an increase in v_s. During cathodic polarization v_s is large, but is small subsequently. The results indicate that v_s varies in a regular way with the surface potential ϕ_s of the germanium, but that chemisorption of hydrogen results in a decrease in the density of surface recombination centers.

INTRODUCTION

THE concepts of traps and recombination centers at semiconductor surfaces remain abstractions. The work reported herein is preliminary to the association of surface energy levels with particular arrangements of surface atoms having specific electronic distributions. The system germanium–aqueous solution was chosen for study because, *inter alia*, it can be made to approach closely the ideal of an uncontaminated (in the electrochemical sense) semiconductor surface in a controlled chemical environment.

As a working hypothesis, it is considered that germanium atoms at a free surface can exist in either of two *stable* configurations, viz. the configuration involving two bonds to the crystal ($>$Ge), or that characterized by three Ge–Ge bonds ($\geq$Ge). (Only a transient existence is ascribed to a singly bonded surface atom, since formation of a second bond to the crystal would necessitate at most a small displacement of the atom from its initial position.) This will be the case, independently of the macroscopic crystallographic orientation of the surface, provided the existence of surface irregularities of atomic dimensions is admitted. The inherent differences in chemical reactivity for the principal orientations of germanium surfaces, {100}>{110}>{111}, as found for oxygenated aqueous solutions by HARVEY and GATOS[1, 2] and for sufficiently dilute etchants by ELLIS[3], can be explained on the basis of different densities of the two kinds of surface atom. It also follows from this work, as pointed out by BECK and GERISCHER[4], that the configuration $>$Ge is the more reactive.

In the absence of an appreciable oxidic layer, it should be possible to observe the effects of direct interactions of germanium surface atoms with aqueous ambients on the electrophysical properties of the germanium. Some results of a preliminary study have been reported.[5] In this work, dissolved oxygen resulted in an increase in surface conductance and in photoconductance and lifetime for germanium filaments in distilled water; the sign and magnitude of the germanium electrode photopotential were also shown to depend in an important way upon the composition of the solution including, in this case, the concentration of electrolyte.

The effects described in the preceding paragraph are interrelated. Thus, as is well known,[6, 7] the surface recombination velocity v_s is a function not only of the nature (distribution in energy, capture probabilities for holes and electrons) and density of surface recombination centers, but also of the equilibrium free-carrier densities in the interior as well as of the surface potential, $\phi_s = \psi_s - \psi_b$, i.e. the drop in electrostatic potential across the region of space charge.

Ideally, one would like to measure v_s for

* Operated with support from the U.S. Army, Navy and Air Force.

VOL. 14 1960

germanium electrodes over a wide range of ϕ_s.* Unfortunately, while it is easy to vary the surface potential of a semiconductor electrode, the actual measurement of ϕ_s (or of changes in ϕ_s) is complicated by the fact that it becomes essential to know, for each value of electrode potential, the fraction of the total germanium–electrolyte potential drop which falls across the space charge region in the germanium. The only direct experimental results bearing on this question were obtained by BOHNENKAMP and ENGELL[8] from measurements of the impedance of the germanium–electrolyte phase boundary. These authors show that a large portion of the voltage *applied* to a germanium electrode is located within the space charge, but their measurements are not quantitative in this respect. There is evidence, moreover, that the densities of surface recombination centers may be altered by electrolytic polarization (see below).

RESULTS AND DISCUSSIONS

In the present paper, some results of a novel approach to the measurement of the variation of v_s with solution composition are described. The method is an adaptation of that employed by PLESKOV[9] for investigating injection and extraction of minority carriers as the result of electrochemical reactions. In electrochemical terms, the system employed for monitoring changes in v_s with solution composition may be represented as in Fig. 1: a germanium wafer separating two solutions,

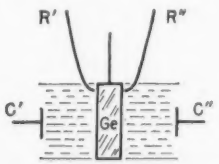


FIG. 1. Diagram showing a germanium disk as the common electrode of two electrochemical cells: R' and R'' are reference electrodes; C' and C'' counter electrodes.

having appropriate reference and counter electrodes, can participate in electrochemical reactions occurring simultaneously at opposite faces. If the wafer is n-type and thin compared to a minority-carrier diffusion length, and if one side is polarized anodically, then the limiting anodic current is a sensitive indication of hole generation or recombination at the opposite surface. With the disks mounted and masked as shown in Fig. 2, very sharp current saturation was obtained, so that

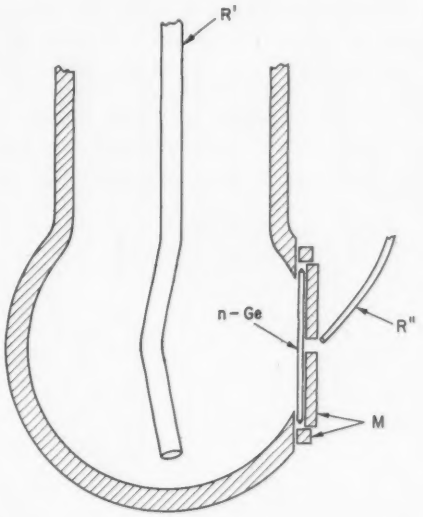


FIG. 2. Apparatus for measuring changes in V_s at a germanium surface in aqueous solutions. The electrical lead to the ohmic ring contact is not shown. Tubing R' leads to solution reservoirs and to a reference electrode for measuring changes in electrode potential at the *back* surface; changes in electrode potential at the *front* surface are measured against a second reference electrode, whose Luggin capillary is designated as R''; parts labeled M are a ring and disk of glass, used in conjunction with Apiezon W wax for masking purposes.

measurement at a single anodic potential (e.g. +1·0 V vs. s.c.e.) was sufficient to establish the magnitude of the limiting current with good precision.

In the region of limiting current, holes are consumed by the anodic reaction as rapidly as they can diffuse to the surface from the bulk and can be supplied by generation processes at the surface. The expression[10] given by equation (1) for the hole saturation current density j_p at an n-type anode applies to an idealized situation in which other surfaces are many diffusion lengths away.

$$j_p = qp_0[(D_p/L_p)+v_s] \qquad (1)$$

In equation (1), p_0, D_p and L_p are the equilibrium

*Quite recently, H. U. HARTEN has carried out such measurements for a silicon electrode (reported at this Conference).

bulk density, the diffusion constant and the diffusion length, respectively, for holes, and q is the charge of a hole. With the thin disk geometry, however, a term must be added to take into account generation at the back surface. Experimentally, a decrease in the magnitude of v_s' at the unilluminated back surface (left side in Fig. 2) is revealed by a decrease in the anodic saturation current crossing the front surface. On the other hand, the anodic limiting current under strong illumination of the back surface is increased as a result of a decrease in v_s'. Because of compensatory effects, the increase in limiting current under weak illumination is related to v_s' and light intensity at the back surface in a rather complicated way. These features are illustrated by typical data presented in Fig. 3.

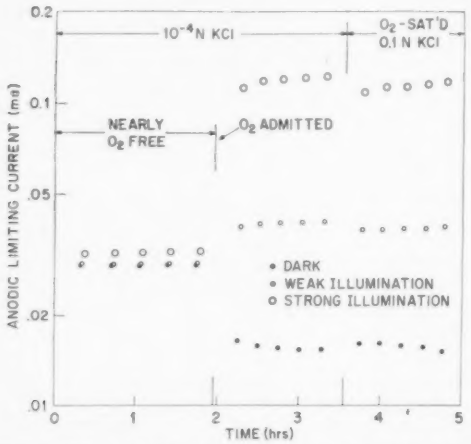


FIG. 3. Effect of illumination and of changing solution composition at the back surface of a thin n-germanium disk (4·6 Ω-cm, {100}, $\tau_b \sim 1300$ μsec) on the limiting current measured during anodic polarization of the front surface (see Fig. 3).

From the foregoing discussion and by reference to Fig. 3, it is evident that the anodic limiting current crossing the front surface is altered in the expected way by interaction of oxygen with the back surface. The effect of dilution of the electrolyte is not nearly so pronounced (found also to be the case in N_2-saturated solutions). More information is needed before the latter results can be correlated with the observed effect of electrolyte concentration on the magnitude of the germanium electrode photopotential.[5] Perhaps the most significant conclusion to be drawn from these measurements is that the reaction of germanium with water and dissolved O_2 is not strongly hole injecting, i.e. the reduction of oxygen does not furnish holes in considerable excess of those required for germanium disolution.

If numerical values of recombination velocity are to be obtained by this method, it is necessary to vary the light intensity in a known way. In order to formulate the problem in simplest terms, it will be assumed that the concentration gradient for holes is constant in the interior, the hole density at the edge of the space charge region at the illuminated back surface having the non-equilibrium value p; the magnitude of p will depend upon the light intensity I and the recombination velocity v_s' at the back surface. With these assumptions, the limiting anodic current density j will be equal to

$$j = \alpha j_p = \alpha q D_p(p/t) + \alpha q p_0 v_s \qquad (2)$$

where α is the current multiplication factor[11] and t is the thickness of the wafer ($t \ll L_p$). Applying the continuity condition

$$I - v_s'(p - p_0) = D_p(p/t) \qquad (3)$$

p may be eliminated, giving the result

$$j = \frac{\alpha q}{1+(t/D_p)v_s'} I + \alpha q p_0 \left[\frac{v_s'}{1+(t/D_p)v_s'} + v_s\right] \qquad (4)$$

Equation (4) predicts a linear variation of the anodic limiting current with light intensity (provided concentration polarization on the electrolyte side of the interface has a negligible effect on the measured current). Further, as v_s' is increased, the slope $\partial j/\partial I$ should decrease and the intercept $j_I = 0$ should *increase*. Qualitatively, this is the behavior observed, as illustrated in Fig. 4. The area A of the germanium–electrolyte contact on the anodic side was 0·026 cm^2 in these measurements. Part of the initial variation of $\partial j/\partial I$ with I may be the result of a decrease in the multiplication factor, i.e. $\alpha = \alpha(p)$.

Fig. 4 typifies results of a study of the effects of cathodic charging on surface recombination velocity. In this study, the anodic limiting current crossing the front surface was measured as a function of light intensity during illumination and cathodic polarization of the back surface. In agreement with the frequently quoted observation by

VOL. 14 1960

ELLIS[12], it was found that the passage of cathodic current can result in a substantial increase in v_s' (compare curve (1) with curves (2) and (4)). It is significant, however, that in no case did the value of v_s' during cathodic charging exceed that obtainable by reducing the concentration of dissolved oxygen to a very low level (cf. curve labeled "N_2", which was obtained on a previous day). Part, at least, of the effect of cathodic charging in solutions not deoxygenated can be attributed to removal of oxygen by electrolytic reduction. Other features to be considered are the following: (a) no change in v_s' was observed when the solution, initially nearly O_2-free by virtue of prolonged bubbling of purified N_2, was saturated with purified H_2, and (b) there is evidence that at higher cathodic current densities (e.g. 50 mA/cm^2), v_s' begins to decrease. This latter observation can be interpreted to indicate that ϕ_s had been shifted to a large negative value, thereby traversing the maximum in the v_s' vs. ϕ_s curve. However, copious hydrogen evolution at the higher cathodic currents may be a disturbing factor in these measurements.

Following cathodic charging (curves (3) and (5)), v_s' was very low; in fact, data obtained on a previous day in O_2-saturated solution coincide almost exactly with curve (5). This behavior is not compatible with the view that discharged hydrogen enters the germanium lattice[9] and there gives rise to a large apparent surface recombination velocity. It should also be mentioned that v_s' is restored to its prepolarization value in times of the order of hours. The mechanism of restoration probably consists in removal of chemisorbed hydrogen by reaction with residual oxygen in the solution.* Since the potential returns to its equilibrium value in times of the order of a minute following cathodic charging, the slow increase in v_s' takes place at essentially constant potential difference between germanium and electrolyte.

It seems probable that the removal of chemisorbed hydrogen results in an increased density of

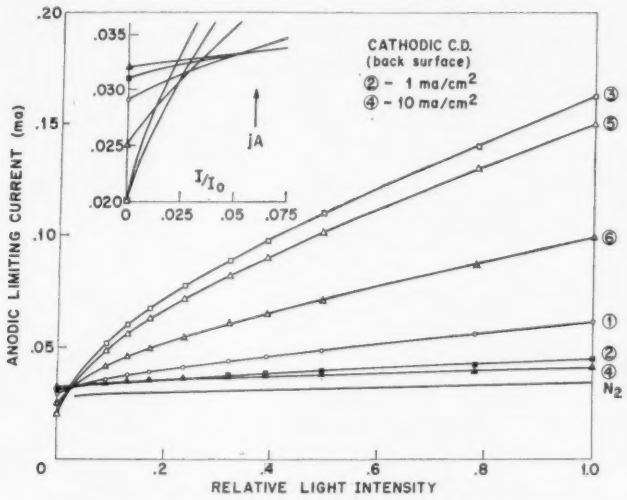


FIG. 4. Effect of cathodic charging in H_2-satd. 0·1N H_2SO_4 on v_s, as indicated by the anodic limiting current at the opposite surface of a thin n-germanium disk; intensity of illumination was varied by means of a calibrated set of neutral density filters. Encircled numbers indicate sequence of measurement; 35 min. elapsed between curves (5) and (6). Inset shows relation between intercept and initial slope.

* At very low concentrations of dissolved oxygen, germanium dissolution rates are of the order of one monolayer per hour or less.[1]

surface recombination centers, which suggests that the centers are associated with the so-called "dangling bonds."

Acknowledgements—The author is pleased to acknowledge the assistance of Mrs. M. J. BUTTON with the experiments and is indebted to Dr. H. U. HARTEN for a helpful discussion of minority carrier distribution in thin semiconducting specimens.

REFERENCES

1. HARVEY W. W. and GATOS H. C., *Trans. Electrochem. Soc.* **105,** 654 (1958).
2. HARVEY W. W. and GATOS H. C., *J. Electrochem. Soc.* **107,** 65 (1960).
3. ELLIS R. C. Jr., *J. Appl. Phys.* **25,** 1497 (1954).
4. BECK F. and GERISCHER H., *Z. Elektrochem.* **63,** 500 (1959).
5. HARVEY W. W. and GATOS H. C., *J. Appl. Phys.* **29,** 1267 (1958).
6. STEVENSON D. T. and KEYES R. J., *Physica* **20,** 1041 (1954).
7. MANY A., HARNIK E. and MARGONINSKI Y., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 85. University of Pennsylvania Press, Philadelphia (1957).
8. BOHNENKAMP K. and ENGELL H. J., *Z. Elektrochem.* **61,** 1184 (1957).
9. PLESKOV IU. V., *Dokl. Akad. Nauk SSSR* **126,** 111 (1959).
10. GARRETT C. G. B. and BRATTAIN W. H., *Phys. Rev.* **99,** 376 (1955).
11. BRATTAIN W. H. and GARRETT C. G. B., *Bell Syst. Tech. J.* **34,** 129 (1955).
12. ELLIS S. G., *J. Appl. Phys.* **28,** 1262 (1957).

VOL
14
196

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 87–93. Printed in Great Britain.

ELECTROKINETIC BEHAVIOR OF GERMANIUM–LIQUID INTERFACES

W. ERIKSEN* and R. CAINES

Raytheon Research Division, Waltham, Massachusetts

VOL. 14 1960

Abstract—When germanium comes into contact with a liquid or condensable vapor through an interface, the possibility exists of a doubly diffuse mobile space charge at the interface. That part of the diffuse double layer on the germanium side has been investigated in inversion layer studies. The part of the diffuse double layer on the liquid side has not been investigated. This paper discusses the properties of that part of the interfacial double layer on the liquid side of the interface.

The methods used to study the charge distribution in the liquid are the classical electrokinetic phenomena with particular emphasis on streaming potential measurements. Liquids are forced through a porous plug of germanium under hydrostatic pressure and the potential created across the germanium is measured.

The results of these experiments are discussed with the aim to provide further clarification of the phenomena of liquid induced excess currents on germanium.

INTRODUCTION

WHEN two substances of different chemical constitutions are placed in contact through an interface, a potential difference is established across the interface. This potential difference can, in general, be divided into two contributions, a surface potential jump X, due to surface polarization and orientation of dipoles, and a contribution D due to displacement of free charges and ions.[1] The total potential difference between the two phases is then

$$\Delta\phi = \chi + D$$

The potential difference between the phases is not measurable and can be approached only by model considerations.[2] However, it is just this potential difference that is important for surface studies on semiconductor liquid or gas systems. In particular, the contribution, D, is the seat of excess conductances which are important in device technology.[3–5] The part D is accessible to experimentation and forms the subject of this paper.

In the particular case of a semiconductor–liquid system the contribution D has two components, one in the semiconductor side of the interface and one in the liquid.[6] That part of D which appears in the semiconductor exhibits itself in many cases as an inversion layer and has been studied extensively. That part of D which appears in the liquid has as yet only been inferred[6, 7] and has received very little attention.

There exists, however, a very well known technique of studying D in the liquid part of an interfacial double layer by electrokinetic phenomena i.e. electro-osmosis or streaming potentials.

These techniques have been applied to the system germanium and several commonly occuring liquid compounds, with a view to understanding more fully the interactions and charge transfers which occur and lead to surface conductances.

CLASSICAL THEORY OF ELECTROKINETIC PHENOMENA

It was found very early in the history of physical chemistry that when a liquid in contact with a solid was sheared by flowing the liquid over the solid an electrical potential was established in the liquid. These effects are called electrokinetic phenomena and have been found to be due to the existence of a mobile charge layer on the liquid side of the liquid solid interface.

There are several ways of deriving quantitative

* Presently at Semiconductor Division.

expressions for electrokinetic effects. Probably the most illuminating is by use of irreversible thermodynamics.[8] ONSAGER's formulation of irreversible thermodynamics states that in a system having several fluxes, J_i, and driving forces, X_i, that

$$J_i = \sum_k L_{ik} X_i \tag{1}$$

where the L_{ik} are phenomenological coefficients.

A second principle is that the product of the rate of entropy production of the system and the temperature is equal to the sum of the products of conjugate forces and fluxes

$$T\frac{dS}{dt} = \sum_i J_i X_i \tag{2}$$

In electrokinetic phenomena we consider a flow of matter and a flow of electric charge.

Consider a tube of length l and cross-section A. Let the electric current density be J and the rate of flow of liquid in volume per unit time be F. The electric potential is E and the pressure difference between the ends of the tube P. Then the rate of entropy production per unit volume can be shown to be[9]:

$$T\frac{dS}{dt} = J\frac{E}{l} + \frac{F}{A}\frac{\Delta P}{l} \tag{3}$$

from which we see that E/l and P/l can be considered driving forces for the fluxes J and F/A.

We can then write

$$J = L_{11}\frac{E}{l} + L_{12}\frac{\Delta P}{l} \tag{4}$$

$$\frac{F}{A} = L_{21}\frac{E}{l} + L_{22}\frac{\Delta P}{l} \tag{5}$$

and apply the Onsager reciprocal relationship $L_{12} = L_{21}$. The physical significance of the L's can be obtained as follows: When there is no pressure gradient equation (4) reduces to

$$J = L_{11}\frac{E}{l} \tag{6}$$

from which it is immediately evident that L_{11} equals the electrical conductivity of the liquid. In the absence of an electric field equation (5) reduces to

$$\frac{F}{A} = L_{22}\frac{\Delta P}{l} \tag{7}$$

from which according to Poiseuille's law

$$L_{22} = \frac{r^2}{8\eta} \tag{8}$$

where r is the radius of the tube and η is the viscosity of the liquid.

The other constant L_{12} (or L_{21}) necessary for a description of electrokinetic phenomena is connected with the strength of the charged layer at the interface.

Before considering the physical significance of L_{12} an expression for streaming potential will be derived from Principles of Irreversible Thermodynamics. Streaming potential is chosen for discussion because of the experimental ease with which it can be measured. It is defined as the electric potential difference per unit pressure difference at stationary state with zero electric current.

$$\frac{E}{P} = \frac{L_{12}}{L_{11}} \quad (J = 0) \tag{9}$$

or

$$E = \frac{PL_{12}}{\lambda} \tag{10}$$

where λ is the electrical conductivity of the solution.

The reasoning from Irreversible Thermodynamics has some strengths. These lie in the complete independence on a detailed model. Experimental data can be correlated in terms of the equations derived. However, herein also lie some weaknesses. No feeling is obtained for the physical significance of $L_{12}(= L_{21})$ and in order to obtain some physical significance a model approach must be invoked.

We use a model which has been proposed previously to explain excess surface conduction on semiconductor surfaces, in which the system semiconductor–liquid is considered as a two phase system with one common component, electrons, and the interface is a membrane permeable to electrons. The condition for thermal equilibrium is that the electrochemical potential (Fermi level)

VOL. 14 1960

for electrons is constant throughout both phases. The attainment of this equilibrium will require charge transfers with a resulting space charge at the interface. That part of the space charge in the liquid is the subject of this discussion. When the liquid is sheared longitudinally to the surface the velocity increases from zero at the surface to μ in the interior of the liquid. Consider a thin layer of liquid near to and parallel to the surface. Let the strength of the charge layer in this strip be $d\tau$ i.e. the electric moment of the strip per unit area. At steady state the force applied to the strip by a field E is just opposed by the frictional force; that is

$$\eta d\mu = E d\tau \tag{11}$$

integrating from the wall to the interior we get with proper boundary conditions

$$\eta\mu = E\tau \tag{12}$$

where τ is the electric moment per unit area of the whole charge layer. The flow of liquid in volume per unit time is

$$F = \mu A \tag{13}$$

where A is a cross-sectional area somewhat smaller than the total cross-section of the tube. However, the extent of the charge layer in the liquid is so small as to be negligible for any capillary tube of reasonable size so A can be considered the total cross-section of the tube. Substituting equation (13) into equation (12) we get

$$\frac{F}{A} = \frac{\tau E}{\eta} \tag{14}$$

while from equation (5) we have that

$$\frac{F}{A} = L_{21}E \tag{15}$$

therefore $L_{21} = L_{12} = \tau/\eta$ and is proportional to the electrical moment of the double layer.

We can substitute for L_{21} into equation (10) to obtain

$$E = \frac{P\tau}{\lambda\eta} \tag{16}$$

In many formulations of electrokinetic effects τ is divided by the rational permittivity and the resulting potential is considered instead of the electric moment

$$\zeta = \frac{4\pi\tau}{\epsilon}$$

This potential is called the *zeta* potential and will be discussed in this paper. In these terms equation (16) becomes

$$E = \frac{PS\epsilon}{4\pi\eta\lambda} \tag{17}$$

where E = electrical potential, P = pressure, η = viscosity of liquid, λ = conductivity of liquid.

From the derivation it follows that the ratio E/P should be a constant and that the streaming potential should be independent of the capillary dimensions. Certain conditions must be satisfied in order to maintain the validity of this equation: (a) the flow must be laminar; (b) the radius of curvature of the capillary or the pores of the plug must be much larger than the thickness of the double layer; and (c) the conduction current should depend solely on the conductivity of the liquid. Conditions (a) and (b) are easily satisfied; however, it is found that for very dilute solutions, which are usually the type under consideration, there is an anomolous conductance which plays an important role.

There is general agreement that the use of the specific conductance *of the liquid "in the bulk" is erroneous.*[10] The conductance of the liquid at the walls of the tube is actually greater than that in the bulk of the liquid and therefore the more correct value for the specific conductance is that of the liquid in contact with the walls of the tube or the pores of the plug. In practice it is found that use of this value compensates for the surface conductance.

EXPERIMENTAL METHODS AND PROCEDURES

In order to compute the ζ-potential from equation (17) it is necessary to measure three variables: (1) the hydrostatic pressure P, (2) the potential across the diaphragm E_s, and (3) the specific conductivity of the liquid in the diaphragm λ. These measurements were made possible by the apparatus diagrammed in Fig. 1.

A plug of ground particles, considered to be a bundle of parallel capillaries, was used rather than a single capillary due to the difficulties involved in machining germanium capillaries. Another factor is the magnification of electrical effects by the greater surface area of the diaphragm.

The liquid contained in F_1 is forced under a known and constant pressure supported by N_2 gas through the manifold to the reference cell C_1 and the sample cell C_2. The height of the liquid column supported in the manometer tube determines the pressure. The flow rate is determined from the volume of liquid flowing per unit time from the sample tube into the calibrated tube.

The specific conductivity of the liquid in the bulk is an important determining factor for the specific conductivity in the diaphragm and the conductivity cell C_1 provides an adequate means of keeping a periodic check on this value.

The potential was recorded directly across the sample by means of a high impedance voltmeter. Platinum wire soldered at the Pt gauze electrodes and extended through the ends of the sample cell were used to make contact through mercury wells with the voltmeter.

The sample cell C_2 was designed to serve as a combination streaming potential–conductivity cell which permitted simultaneous measurements of the potential and specific resistance of the liquid in the diaphragm.

The sample of germanium particles was confined in the area between the platinum electrodes of the sample cell. The size of the particles was restricted by the size of the platinum mesh electrodes. The absolute size is not known; however, a relative determination was made. The particles were sifted from a 40-mesh to a 60-mesh sieve and those retained in the 60-mesh sieve were selected for use. These were graded to be 0·0165–0·00098 in. or 0·42–0·25 mm in diameter as determined from the sieve openings. Only one packing was used for the various measurements.

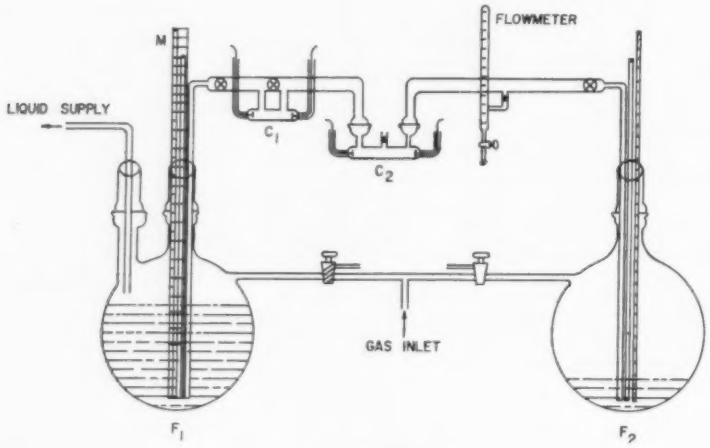


FIG. 1. Apparatus.

The procedure followed was first to exert a pressure on the liquid and adjust until a constant pressure reading was observed. The liquid flowed under this pressure through the sample cell until steady state as indicated by a constant potential reading, was attained. The cell potential was always recorded before streaming was begun. In order to determine the rate of flow, the stopcocks to the calibrated tube at the end of the manifold was opened to the streaming liquid, following the potential readings at each pressure. An immediate drop in pressure was observed, due to the increase in volume; however, when the original pressure was regained the number of cm^3/sec flowing were recorded. An electric stopclock was used for timing at 30-sec intervals. The streaming potential was observed at different pressures from about 10–100 cm of liquid pressure.

VOL. 14 1960

Following the potential measurements, the resistance of the sample cell containing the liquid in contact with the diaphragm was measured. The "cell constant" for the conductivity and sample cells was determined in the conventional way. The resistance r of the cells containing standard solutions of known specific conductivity λ were measured and from the equation $L_r = K$, the cell constant K was calculated. This value could then be used to determine the specific conductivity of any liquid contained in the cell for which the resistance is known. For the Ge–H_2O system, a 0·02 N KCl solution, specific conductivity 0·002760 $\Omega^{-1}cm^{-1}$, was used and for the other three systems a 0·10 N KCl solution, $\lambda = 0{\cdot}011856\ \Omega^{-1}cm^{-1}$. In Table 1 the data from these measurements are recorded.

Table 1. Specific conductivity (λ) data

System	Specific conductance (Liq. in diaphragm) (mho)	Specific conductance (Liq. in bulk) (mho)
Ge–H_2O	$6{\cdot}09\times10^{-6}$	$1{\cdot}5\times10^{-6}$
Ge–CH_3OH	6×10^{-5}	$1{\cdot}4\times10^{-6}$
Ge–CH_3COCH_3	$2{\cdot}96\times10^{-5}$	$4{\cdot}4\times10^{-7}$
Ge–$C_6H_5NO_2$	$7{\cdot}7\times10^{-7}$	$1{\cdot}19\times10^{-8}$

Germanium–conductance-water system

Conductance water having a specific conductivity of $2\times10^{-6}\ (\Omega\ cm)^{-1}$ was used in the determination of the streaming potential. Stock distilled water was re-distilled in a 20-plate column and distillation was controlled at a reflux ratio of 20:5.

The range of pressures exerted gave rise to streaming potentials varying from approximately 0·050 to 0·25 V. The results obtained, when the specific conductivity of the water in the sample cell was used, are given in Table 2. The average value for ζ-potential is 0·02 V.

Germanium–methanol, –acetone, –nitrobenzene systems

These systems can be treated collectively since the only variable in the experimental set-up was the liquid itself. The same cell and packing was used throughout. Reagent grade stock solutions of these organic liquids were used.

Table 2. Streaming potential data and ζ potential for Ge–H_2O System

E_s (V)	P (mm of H_2O)	E_s/P	ζ (V)
0·06	27·8	0·0022	0·019
0·105	47·2	0·0022	0·020
0·172	67·3	0·0025	0·023
0·20	86·5	0·0023	0·020
0·225	107	0·0021	0·019

The entire system was thoroughly cleaned and rinsed with distilled water. The liquid in question was then allowed to flow through the re-assembled system until a constant reading for the resistance of both the conductivity and sample cells were observed.

For these systems, potentials could not be observed below liquid pressures of approximately 60 cm, a range of from 35 to 55 mm Hg. The pressures in these cases are therefore from about 60 to 105 cm-liquid pressures. The methanol system exhibited potentials ranging from about 15 to 30 mV, the acetone systems from approximately 180 to 480 mV, and nitrobenzene from 12 to 20 mV. The significance of these values will be treated in the discussion section.

An average of the three streaming potential runs for the zeta potential determination for each system was taken for the final value. The streaming potential data for these runs are given in Tables 3, 4 and 5.

Table 3. Streaming potential data and ζ potential for the Ge–CH_3OH system

P (cm–CH_3OH)	E_s (V)	E_s/P	ζ (V)
76·5	0·0080	0·000104	0·025
86·5	0·015	0·000173	0·025
96·5	0·023	0·000238	0·035

Table 4. Streaming potential data and ζ potential for the Ge–CH_3COCH_3 *system*

P (cm–CH_3COCH_3)	E_s (V)	E_s/P	ζ (V)
71·5	0·183	0·0034	0·22
76·5	0·217	0·0030	0·19
81·5	0·262	0·0032	0·21
86·5	0·358	0·0031	0·20
96·5	0·479	0·0036	0·23

Table 5. Streaming potential data and ζ-potential for the Ge–$C_6H_5NO_2$ *system*

P (cm–$C_6H_5NO_2$)	E_s (V)	E_s/P	ζ (V)
61·5	0·0122	0·877	0·00020
71·5	0·0135	0·815	0·00019
81·5	0·0155	0·795	0·00019
91·5	0·0171	0·802	0·00019
101·5	0·0194	0·830	0·00019

DISCUSSION

For the particular liquids investigated in this study a positive space charge is accumulated in the semiconductor surface and a negative space charge in the liquid. There is an additional potential jump at the surface, the X-potential, which is a function of the surface polarization and dipole layers existing at the interface. The sum of these three potentials is related to the single electrode potential for the system. Unfortunately, the single electrode potential is not measurable, in fact is not even thermodynamically definable, so that it appears to be impossible to carry through a quantitative description of the potential through an interface from one phase to another. A definition can only be applied to the two space charge regions of different polarity which go to make up the mobile parts of the interfacial double layer. We can define a surface potential in the semiconductor referred to a zero of potential in the bulk of the semiconductor and a surface potential in the liquid referred to a zero of potential in the bulk of the liquid.

It will be fruitful to consider for a moment the parallels between the results reported here on the diffuse charge layer in the liquid, and studies on inversion layer which is the diffuse charge layer in the semiconductor. Together these make up the double interfacial charge layer, D. One of the experimental results obtained from inversion layer studies is the surface potential, ϕs. In the case discussed here we obtain a potential which is the correlary to ϕs on the liquid side of the double layer. The experimental approaches to the respective evaluations of the surface potentials are quite different. In the inversion layer case the evaluation depends on measurement of the conductivity of the inversion layer and a detailed knowledge of the conduction mechanism of the semiconductor. In the present investigation the evaluation of ζ depends on a phenomenological formulation with only secondary dependence on a model and no dependence on the conduction mechanism in the liquid. This lack of knowledge of the mechanism of conduction in the liquid imposes certain restrictions on the information that can be obtained from the experiments.

Ideally, the information wanted from the experments performed here is the charge distribution in the diffuse layer in the liquid. The experimental result obtained is essentially the electric moment of the double layer which is expressed as a potential by dividing by the rational permittivity. The separation of the electric moment into its components of charge and coordinate cannot be accomplished without further knowledge of the conduction mechanism in the liquid. This information is lacking so one is forced to discuss the results in terms of τ, the electric moment or ζ, the zeta potential.

The interesting result from this investigation is that the organic liquids acetone, methanol and nitrobenzene appear to be as active electrically in contact with a germanium surface as water. In fact acetone is much more active judging by the large value of zeta potential for the acetone–germanium interface. It is very difficult to understand how this can occur by invoking the classical theories of selective adsorption of ions at the solid surface. For this reason we have considered the space charge in the liquid to be the result of electron transfer between liquid and semiconductor.

REFERENCES

1. Adams N. K., *Physics and Chemistry of Surfaces*. Oxford University Press (1941).
2. Guggenheim E. A., *Thermodynamics*. North-Holland Publishing Co., Amsterdam (1957).

3. STATZ H., DE MARS G., DAVIS L. and ADAMS A., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 139. University of Pennsylvania Press, Philadelphia (1957).
4. BROWN W. L., *Phys. Rev.* **91,** 518 (1953).
5. KINGSTON R. H., *Phys. Rev.* **93,** 346 (1954).
6. ERIKSEN W. T., *J. Appl. Phys.* **29,** 730 (1958).
7. LAW J. T., *Proc. Inst. Radio Engrs., N.Y.* **42,** 1367 (1954).
8. ONSAGER L., *Phys. Rev.* **37**, 405 (1931).
9. DEGROOT W., *Thermodynamics of Irreversible Processes.* Interscience, New York (1951).
10. KRUYT H. R., *Colloid Science,* Elsevier Publishing Co., Amsterdam (1952).

VOL.
14
1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 94–103. Printed in Great Britain.

DISPLACEMENT PROCESSES ON GERMANIUM SURFACES

K. H. MAXWELL and MINO GREEN

Zenith Radio Corporation, Chicago 39, Illinois

Abstract—Some displacement experiments were carried out on germanium surfaces produced by crushing. In these a monolayer of either water, methanol, n-butanol, isopropanol or atomic hydrogen was first adsorbed, the excess gas removed from the gaseous phase and then a known amount of oxygen admitted to the system. The results show that oxygen reacted with the germanium surface in spite of the material already adsorbed thereon.

For those surfaces covered with a monolayer of water it was found that oxygen displaced most of the water from the surface. The take-up of oxygen was retarded in the early stages of the reaction, compared with a clean germanium surface, but was greater at longer times. The logarithmic rate law was *not* obeyed. Where alcohols were the "pre-adsorbed" material the results were qualitatively very similar to water but the retarding effect at the beginning of oxygen adsorption was greater for the larger molecules.

Oxygen adsorbed to a considerable extent on surfaces covered with a monolayer of hydrogen and, as with the other materials, the early take-up of oxygen was less than with a clean surface. The oxygen take-up after long times was approximately the same as that on the clean germanium surface. Only a small fraction, if any, of the hydrogen was desorbed during the adsorption of oxygen. This certainly suggests that the binding of the hydrogen atoms to the germanium surface is not of the simple covalent type that might have been expected.

INTRODUCTION

OF prime importance to an understanding of any adsorption process is a knowledge of the type of bond holding the adsorbate to the adsorbent. This information becomes doubly important when the adsorbent is a solid semiconductor since an adsorbed layer can affect the surface space charge and its attendant properties considerably.[1] In many instances the bonding is clearly of a particular type but certain cases have been reported recently,[2, 3] involving germanium as the adsorbent, in which it is not certain what bond type was operative. Experiments providing further information about these adsorptions would be useful. In particular, it seemed to the authors that displacement reactions, in which one chemical species replaces another already bound to the germanium surface, would prove informative.

PREVIOUS WORK

In a recent paper the authors[2] reported that the adsorption of certain chemical species on clean germanium surfaces at room temperature occurs to the extent of about a monolayer. Adsorption was very rapid and the rates could not be measured. These adsorbate molecules all contained a hydrogen atom linked to a highly electronegative atom, either oxygen or nitrogen, i.e. the molecules all contained $-OH$ or $>NH$ groups. On the basis of a crude model, electrostatic interaction between the permanent charge distributions of adsorbate and adsorbent was found to account for adsorption. The binding was described as a form of "hydrogen bond". The calculated heats of adsorption were large for molecules which can form a hydrogen bond (water, methanol, etc.) and virtually zero for molecules not of this type (chloroform, benzene, etc.). The success of the model in predicting the adsorption or non-adsorption of a variety of different molecular types was striking. A consequence of this picture is that the adsorbate molecule is bound closely to a germanium surface atom and the concept of an adsorption site is retained.

The adsorption of molecular hydrogen on a germanium surface at room temperature occurs

to a very slight extent, if at all.[2, 4] However, atomic hydrogen does adsorb strongly[5–7] and the coverage may reach a monolayer. It would be expected that a hydrogen atom would react with the "dangling bond" of a germanium surface atom and a normal covalent bond would ensue. However, HEILAND and HANDLER[3] found that the adsorption of atomic hydrogen increased the *p*-type conductivity of a germanium surface, already strongly *p*-type in the cleaned condition. Had a normal covalent bond been formed a reduction in the *p*-type conductivity would have been expected.

From this brief survey it can be seen that the chemisorption of oxygen does seem to occupy a unique position in the surface chemistry of germanium. The extensive take-up of oxygen beyond monolayer coverage[4, 8] is exhibited by no other adsorbate so far studied. The very large heat of adsorption[8, 9] is evidence that the binding of the initial oxygen layer to the germanium must be covalent.

VOL. 14 1960

SCOPE OF THE PRESENT INVESTIGATION

Some displacement reactions were carried out to obtain more information about the binding of hydrogen, water and some alcohols on germanium surfaces. Hydrogen, water or an alcohol (methanol, isopropanol and n-butanol were employed) was adsorbed on the germanium powder surface to approximately monolayer coverage, the excess gas removed from the gas phase and then a known amount of oxygen admitted to the system. The oxygen pressure was measured as a function of time. Surface areas of the powders were determined. The results are discussed in terms of the binding of the primary adsorbate to the germanium and its effect on the kinetics and extent of oxygen take-up.

EXPERIMENTAL METHODS

(a) *General*

The general experimental details regarding the vacuum system, pressure measurement, purity and purification of materials have been given elsewhere.[2, 8]

(b) *Adsorption apparatus*

Two types of adsorption cells were used, both being a modification of the double chamber apparatus described by GREEN *et al.*[8]. Fig. 1(a) is a diagram of the adsorption cell used whenever the primary adsorbate was a readily condensable gas such as water or an alcohol. The apparatus employed when the first adsorbate was atomic hydrogen is shown diagrammatically in Fig. 1(b).

(c) *Experimental procedure*

The apparatus was evacuated and baked out at 300°C for several hours at 10^{-6} mm Hg. A known quantity of oxygen was admitted to the upper chamber which was then sealed off. The remaining steps in the procedure depended on which gas was the primary adsorbate.

When water or one of the alcohols was the first adsorbate the adsorption cell shown in Fig. 1(a) was used. A known pressure of gas was admitted to the lower chamber which was then sealed off at the point *s*. The germanium disc was crushed and the pressure decrease measured. The excess gas was condensed in the capillary side-arm by placing a liquid nitrogen trap around the side-arm which was kept chilled throughout the experiment. To start an experiment the break-seal connecting the upper and lower compartments was broken to admit the oxygen. The pressure was measured as a function of time.

The adsorption cell shown in Fig. 1(b) was employed for the experiments where atomic hydrogen was "pre-adsorbed" on the germanium surface. A known quantity of hydrogen was introduced into the lower chamber and the point *y* was sealed. The heating current for the tungsten filament was turned on and the germanium crushed. (The conditions under which this operation was carried out will be described in detail later.) When the adsorption of atomic hydrogen was considered to be complete, a break-seal leading to a Varian VacIon pump, working at 10^{-8} mm Hg, was broken and the excess hydrogen pumped off. The adsorption cell was sealed off at the point *z* when the pressure had dropped to approximately 10^{-7} mm Hg; this took typically 2 min. The break-seal connecting the upper and lower compartments was broken to admit oxygen and start the experiment. The pressure was measured as a function of time. All adsorptions were at 23° ± 1°C.

(d) *Supplementary experiments*

Certain additional experiments were undertaken to furnish experimental documentation of certain

aspects of the work. These are described in this section.

It is important to establish, for the experiments using apparatus, Fig. 1(a), that water or an alcohol will be frozen out in the capillary side-arm a short time after entering the gas phase. A mixture of oxygen and water was introduced into a cell that consisted essentially of the bottom compartment of apparatus, Fig. 1(a), and a liquid nitrogen dewar was placed around the side-arm. One min. after cooling the side-arm the thermistor registered a pressure of 135 μ Hg, while after 3 min. the pressure was 132 μ Hg of oxygen, at which value it remained for as long as the measurements were made (3 hr). Thus it can be seen that water condensed out rapidly, even in the presence of the oxygen, and that the pressures recorded by the thermistor will indeed be due to oxygen except possibly for the earliest measurements (the first few minutes). Similar experiments carried out using methanol or n-butanol mixed with oxygen gave similar results, although the condensation of n-butanol was appreciably slower.

A potential source of error was that water, orginally adsorbed on the glass walls of the cell may desorb when oxygen was introduced into the lower chamber. A blank experiment was run in a simple double chamber apparatus[8] having a known amount of oxygen sealed in the upper chamber. The lower chamber was exposed to water vapour at 500 μ Hg for several hours, the water was quickly pumped out and the chamber sealed off. After breaking the connecting break-seal the oxygen pressure was 98 μ Hg and it remained stable over 18 hr. The expected value, knowing the oxygen pressure in the upper chamber and the volumes, was 97·6 μ Hg. It is clear that oxygen does not displace the water adsorbed on the glass, at least to within the accuracy of the pressure measurements (2 per cent).

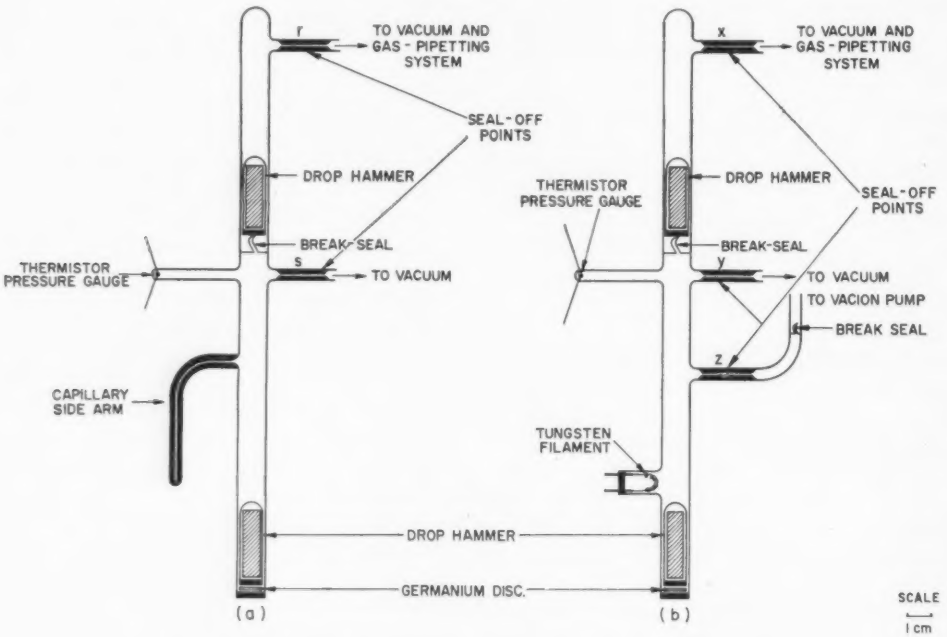


FIG. 1. Diagrams of the two types of pyrex glass adsorption cells used: (a) for the condensable primary adsorbates water and the alcohols; (b) for atomic hydrogen as primary adsorbate.

(e) *The adsorption of atomic hydrogen*

The production of atomic hydrogen was accomplished by a hot tungsten filament[10, 11] using the apparatus shown in Fig. 1(b). The hammer in the lower chamber was raised to a position above the side-arm containing the tungsten filament and the filament heating current was turned on. There was an immediate decline in the pressure which continued for some time but had virtually ceased after 20 min. The germanium was crushed with the glass-encased drop-hammer, the time of crushing being about 3 min., and the hammer was raised to its former position. The pressure had dropped by a large amount and the pressure decline was followed for about 30 min. after which time it had virtually halted. The germanium powder was stirred frequently by applying a vibrator to the bottom of the lower chamber. The results obtained in an actual experiment are shown in Fig. 2; the surface area of the germanium powder was 301 cm^2 and the volume of the lower chamber, plus the tubing up to the break-seal leading to the VacIon pump,

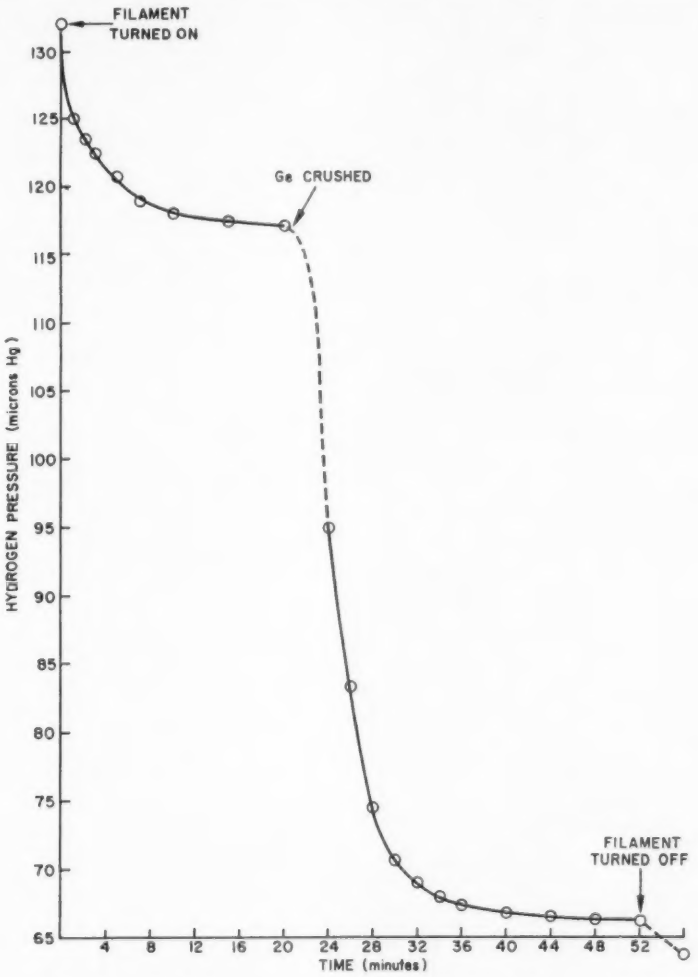


FIG. 2. The hydrogen pressure in the system as a function of time during the adsorption of atomic hydrogen on germanium.

was 84·3 cm³. The decrease in pressure in the time between turning on the filament current and crushing the germanium must be due to the adsorption of atomic hydrogen on the glass; the hydrogen per germanium surface atom, assuming there are $7{\cdot}7 \times 10^{14}$ sites per cm² on a germanium surface.[8] While the adsorption of atomic hydrogen on the glass does prevent an accurate estimate

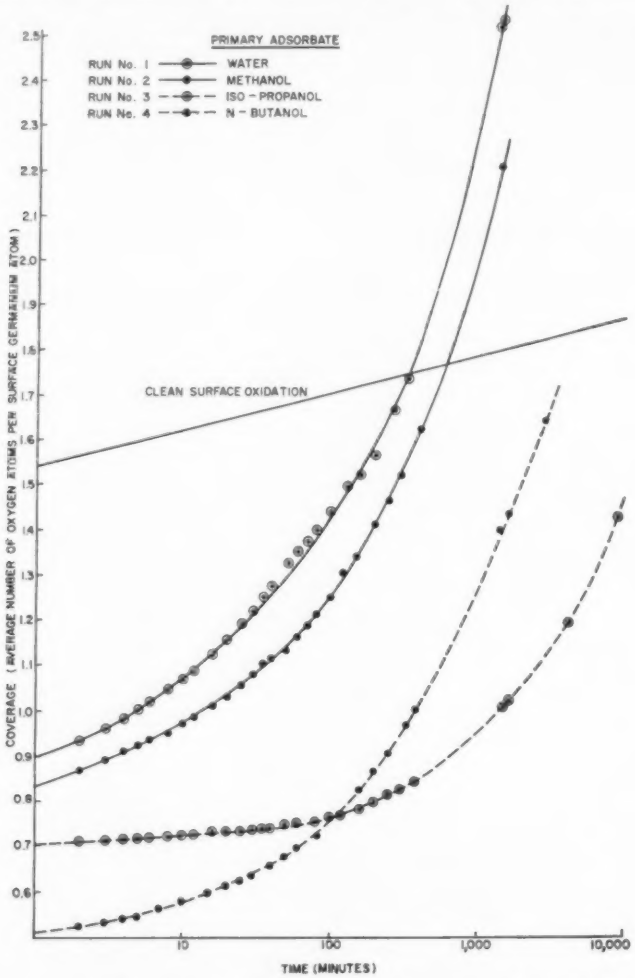


FIG. 3. The take-up of oxygen plotted against the logarithm of the time for germanium surfaces previously covered with a monolayer of water, methanol, isopropanol and n-butanol. Run no. 1 (final pressure) oxygen 93 μ Hg, surface area 198 cm².; Run no. 2 (112 μ Hg, 219 cm²); Run no. 3 (108 μ Hg, 189 cm²); Run no. 4 (89 μ Hg, 197 cm²).

amount of hydrogen so removed was $8{\cdot}2 \times 10^{16}$ *atoms*. The quantity of hydrogen removed from the gas phase after crushing the germanium was about $2{\cdot}8 \times 10^{17}$ *atoms*, or about 1·2 atoms of being made of the quantity adsorbed on the germanium, the results leave little doubt that the coverage was extensive and approached a monolayer.

A blank experiment was carried out to determine whether the hydrogen adsorbed on the glass was desorbed when oxygen was admitted. A similar apparatus to that shown in Fig. 1(b) was used except that neither the germanium nor the hammer was present. The hydrogen was admitted to the lower chamber and the filament turned on for 30 min. The excess hydrogen was pumped out, the lower chamber sealed off, and the connecting break-seal broken to admit the oxygen. The anticipated pressure was 158 μ Hg while the measured pressure was 156 μ Hg and it stayed constant over 24 hr. Thus oxygen does not displace the adsorbed hydrogen from the glass to within the accuracy of the pressure measurements (2 per cent).

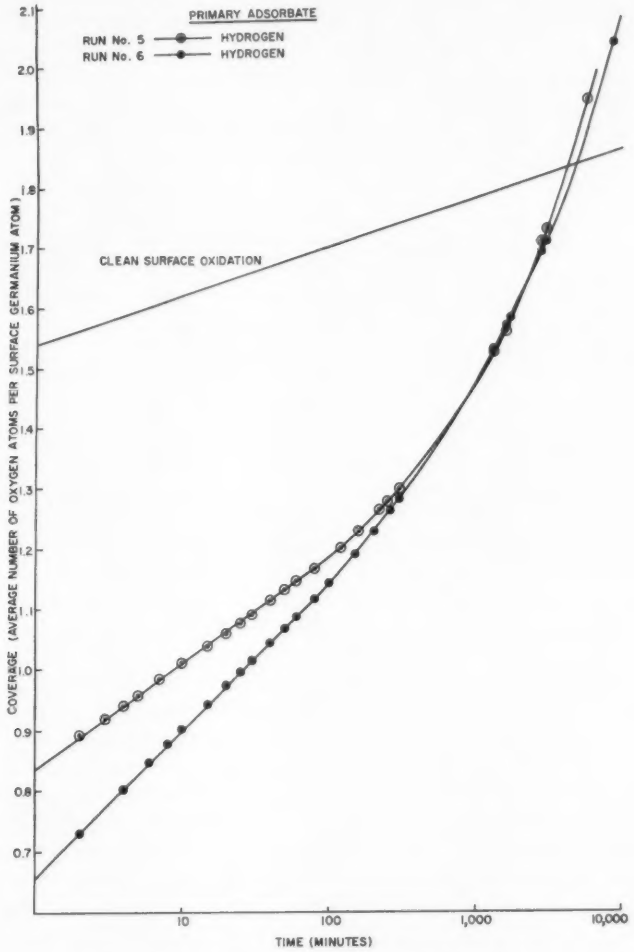


FIG. 4. The take-up of oxygen plotted against the logarithm of the time for germanium surfaces previously covered with a monolayer of hydrogen. Run no. 5 (final pressure oxygen 138 μ Hg, surface area 301 cm^2); Run no. 6 (33 μ Hg, 310 cm^2).

VOL. 14 1960

(f) *Surface area determinations*

When either water or methanol was the primary adsorbate the surface area was calculated from the pressure decrease on crushing and the volume of the chamber. The authors have shown[2] that these two gases adsorb to monolayer coverage at room temperature and at pressures of 50–300 μ Hg. When n-butanol, isopropanol or hydrogen was the pre-adsorbed species the surface area was determined at the end of each experiment by the BET method, using krypton at liquid nitrogen temperature.[12]

(g) *Accuracy*

The take-up of oxygen could be determined with an accuracy of ± 2 per cent. The surface area determinations were reproducible to about ± 3 per cent. The absolute accuracy of the BET method is uncertain[12, 13] but a reasonable estimate would seem to be ± 10 per cent.

RESULTS

The results are summarized in graphical form in Figs. 3 and 4 in which the coverage (the average number of oxygen atoms per germanium surface atom) is plotted against the logarithm of the time. The coverage at a particular time was calculated from the oxygen take-up and the surface area of the powder, it being assumed that the pressure registered by the thermistor was due only to oxygen. The thermistor is a non-linear Pirani gauge which has to be calibrated for each gas and it cannot be used for gas mixtures. A further assumption needed is a figure for the number of adsorption sites per cm^2 of germanium surface. The figure of $7{\cdot}7 \times 10^{14}$ sites/cm^2 was chosen.[8] The pressure in the system at the time the experiment was discontinued and the surface area of the germanium powder for each run are given in the figure captions. At least two runs were carried out with each primary adsorbate, although the results of only one run for each substance are given in Fig. 3. The pressures of oxygen are roughly the same for all the runs shown in Fig. 3. No attempt was made to investigate the effect of oxygen pressure on the rates. For the purposes of comparison the take-up of oxygen on a clean germanium surface (crushed powder) at about the same oxygen pressure is included in Figs. 3 and 4.

DISCUSSION

The results leave little doubt that oxygen reacted with the germanium surface to a considerable extent in spite of the material pre-adsorbed thereon. It is most convenient to divide the discussion into two parts, one dealing with the surfaces on which water and the alcohols were primary adsorbates (Fig. 3) and the other with the surfaces covered by hydrogen (Fig. 4.).

When a germanium surface, covered with a monolayer of water, was exposed to oxygen the adsorption of oxygen was retarded in the early stages compared with that on a clean surface. However, the take-up after 500 min was already considerably greater than for the clean surface and was proceeding at a greater rate. The presence of the initial water monolayer has quite altered the kinetics of the oxygen adsorption and probably the structure of the oxide layer. The coverage after 1400 min exposure to the oxygen was 2·5, already much greater than that ever observed at these pressures when the starting surface was clean germanium.

It is important to establish whether the initial monolayer of water was completely displaced by the oxygen or whether all, or part of it, was retained on the surface during oxygen adsorption. At the end of the oxidation run the condensed water was released by removing the liquid nitrogen dewar from the capillary side-arm and the new pressure measured. (The thermistor usually cannot be used for measuring the pressure of gas mixtures, but the calibrations for oxygen and water vapour were very similar and, in this case, an estimate could be made.) Knowing the oxygen pressure, the amount of water originally adsorbed on the germanium and the excess water frozen out in the side-arm, it was found that approximately 85 per cent of the adsorbed water layer had been displaced. It is difficult to estimate the error in this percentage but ± 10 per cent would be an upper limit. An additional difficulty encountered in making this estimate was that, after the release of the water, the pressure rose rapidly to a maximum and then declined continuously. The fall in pressure was followed for about an hour. No quantitative data could be obtained, for the reasons given above, but it seems most likely that the adsorption of water on the oxidized surface was occurring. On this assumption the amount adsorbed in 1 hr

VOL. 14 1960

corresponded very approximately to 0·4 of a monolayer.

The results obtained for methanol or n-butanol as the primary adsorbate were qualitatively similar to those with water (see Fig. 3). The early take-up of oxygen was considerably less than that on clean germanium but the rates of take-up were much greater at longer times. The surfaces covered with water, methanol and n-butanol seem to form a homologous series in their reaction with oxygen, the larger the grouping attached to the $-OH$ the slower the reaction with oxygen. It was not possible to determine how much of the initial adsorbed alcohol was retained by the oxide layer. There was evidence, similar to that obtained with water, that methanol adsorbed on the oxidized germanium surface.

The surface initially covered with isopropanol behaved rather differently. The initial take-up of oxygen was greater than for the surface covered with n-butanol but the subsequent take-up was much slower. The shape of the curve (Fig. 3) is different from the other three. As this was the only branched-chain alcohol studied, the result suggests that there is a structural factor, other than the mere size of the pre-adsorbed molecule, involved in the retardation of oxygen take-up.

These results are consistent with the authors' view that molecules having the $-OH$ grouping are bound to a germanium surface by electrostatic forces[2], i.e. by forces weak compared with a germanium–oxygen electron pair bond. Had the binding of water and the alcohols been by means of a strong electron pair bond it is unlikely that adsorption of oxygen on a vacant surface germanium site would have induced the desorption of the first adsorbate species. If oxygen adsorption does not seriously perturb the binding energy of the first adsorbate species then the rate of oxygen adsorption should equal the desorption rate of the first species. The desorption rate of water and methanol from a germanium surface at room temperature is known to be small.[14, 15] Therefore, the adsorption of oxygen cannot be simply by the reaction at sites vacated by the unassisted desorption of alcohol molecules.

The following picture of the displacement of water presents itself. Oxygen adsorbs initially in vacant sites left in the adsorbed monolayer of water (for example, 5 per cent of vacant sites could not be detected). The adsorbed oxygen interacts with the neighboring suface complex ($Ge \ldots H_2O$), in such a way that the complex becomes unstable and water molecules are desorbed. More oxygen then adsorbs in the newly created vacant sites and the process continues. This picture visualizes the displacement process as starting at vacant adsorption sites and continues with the oxygen "peeling off" the water layer. On patches of the surface where a monolayer of oxygen atoms has been laid down, the build-up of the oxide layer can commence. The overall kinetics would be very complicated and the rate of oxygen take-up probably sensitive to surface dislocations, crystal edges and corners, these being obvious locations for vacant sites in the monolayer of adsorbed material.

The results obtained when oxygen reacted with a germanium surface covered with a monolayer of hydrogen (Fig. 4) have certain features in common with those just discussed. The early take-up of oxygen is less than that on a clean surface but the *rate* of adsorption is greater and the coverage does eventually exceed that found on the clean surface. The results in Fig. 4 are based on the assumption that hydrogen was not present in the gas phase and the thermistor was reading oxygen pressures. If hydrogen had been displaced from the surface and replaced oxygen in the gas phase the thermistor would have registered an apparent pressure increase. Hydrogen is a more effective conductor of heat than oxygen and hence a given pressure of hydrogen will produce a larger change in the resistance of the thermistor than the same pressure of oxygen. This is illustrated in Fig. 5 which shows typical calibration curves for the two gases. The desorption of small amounts of hydrogen ($<0{\cdot}1$ of a monolayer) could not have been detected. Thus, it can be said that the considerable adsorption of oxygen, measured from the earliest stages of the reaction, was not accompanied by the displacement of large amounts of hydrogen.

The possibility that there was a reaction between oxygen and hydrogen at the surface cannot be ignored. To test this possibility a run was carried out with an apparatus of the type shown in Fig. 1(b) but with an additional capillary side-arm sealed onto the lower chamber. The adsorption of atomic hydrogen was carried out in the manner described previously. Just before admitting oxygen the side-arm was cooled to liquid nitrogen

temperature. Twenty minutes after the admission of oxygen the liquid nitrogen dewar was removed and the side-arm allowed to warm up. No increase in the pressure of the system was noted. Had a significant amount of water been formed and desorbed, it would have been observed subsequently as a pressure rise when the side-arm warmed up. Therefore, it can be concluded that appreciable amounts of water were not formed when oxygen adsorbed on a germanium surface already covered with a monolayer of hydrogen.

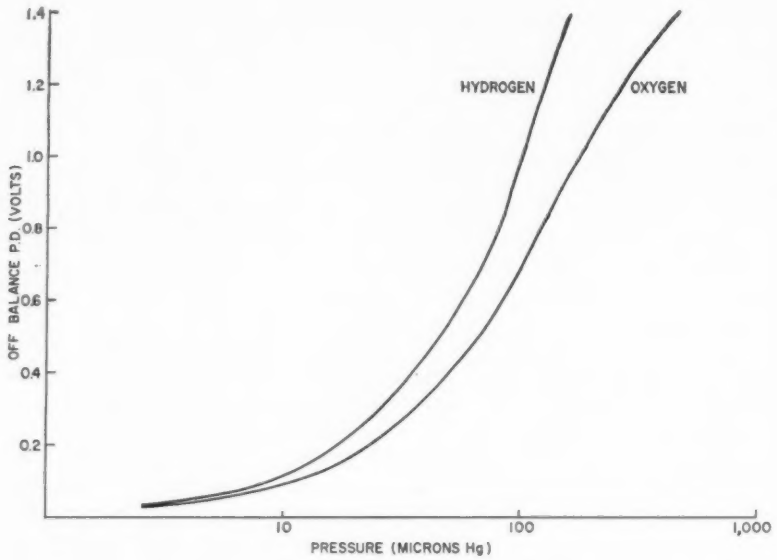


FIG. 5. Calibration of the thermistor pressure gauge for hydrogen and oxygen. The off-balance p.d. is plotted against the logarithm of the pressure.

Since the hydrogen does not leave the surface when oxygen is adsorbed, it must either remain on the surface in some form or else enter the germanium lattice. In this latter connection it must be said that the large heat of adsorption of oxygen[8, 9] could lead to localized surface heating, thereby assisting the dissolution of hydrogen atoms in the germanium. The view of HEILAND and HANDLER[3] that the binding of hydrogen to a germanium surface is complex and not a simple bond would seem to be well substantiated by the results of this work. The binding of the initial hydrogen layer cannot be covalent, the same reasoning applying here as to the water case discussed above. It is tempting to postulate that the hydrogen atoms are bound in positions *between* the germanium surface atoms but there is no direct evidence to support this idea.

SUMMARY

The following conclusions can be drawn from this work concerning the surface chemistry of germanium:

(1) Oxygen adsorbs on germanium surfaces already covered with a monolayer of water or a simple alcohol. For those surfaces covered initially with water or methanol the extent of oxygen adsorption is considerably greater than that observed on clean surfaces.

(2) At least for the water case, most of the initial monolayer is displaced by the oxygen.

(3) Water adsorbs strongly on a clean germanium surface but does not adsorb appreciably on the surface in the *early* stages of oxidation (i.e. oxygen coverage up to approximately a monolayer). During the later stages of oxidation (oxygen

coverage greater than two) water adsorbs by what appears to be a strong physisorption or a weak chemisorption.

(4) These results are consistent with the authors' view that water and alcohols adsorb on a clean germanium surface by a type of hydrogen bond and are not covalently bound.

(5) Oxygen adsorbs strongly on surfaces already covered with a monolayer of hydrogen. The hydrogen is not displaced, either as molecular hydrogen or in the form of water.

(6) The suggestion of earlier workers that the binding of hydrogen atoms to the germanium surface is not by a simple covalent bond appears to be substantiated by the present work.

VOL. 14 1960

REFERENCES

1. BARDEEN J., *Phys. Rev.* **71,** 717 (1947).
   HANDLER P., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 23. University of Pennsylvania Press, Philadelphia (1957).
2. GREEN M. and MAXWELL K. H., *J. Phys. Chem. Solids* **11,** 195 (1959).
3. HEILAND G. and HANDLER P., *J. Appl. Phys.* **30,** 446 (1959).
4. DELL R. M., *J. Phys. Chem.* **61,** 1584 (1957).
5. LAW J. T., *J. Phys. Chem.* **59,** 543 (1955). This reference and Ref. (6) showed the adsorption of atomic hydrogen, but the fact was not recognized at the time. See Ref. (7).
6. SCHLIER R. E. and FARNSWORTH H. E., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 3. University of Pennsylvania Press, Philadelphia (1957).
7. LAW J. T., *J. Chem. Phys.* **28,** 511 (1958).
8. GREEN M., KAFALAS J. A. and ROBINSON P. H., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 349. University of Pennsylvania Press, Philadelphia (1957).
9. BRENNAN D., private communication.
10. LANGMUIR I., *J. Amer. Chem. Soc.* **37,** 417 (1915).
11. LAW J. T., *J. Chem. Phys.* **30,** 1568 (1959).
12. ROSENBERG A. J., *J. Amer. Chem. Soc.* **78,** 2929 (1956).
13. JACOBS P. W. M. and TOMPKINS F. C., *Chemistry of the Solid State* (edited by GARNER W. E.) p. 91. Academic Press, New York (1955).
14. HANDLER P. and PORTNOY W. M., *Phys. Rev.* **116,** 516 (1959).
15. LAW J. T. and FRANCOIS E. E., *Ann. N.Y. Acad. Sci.* **58,** 925 (1954).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 104–110. Printed in Great Britain.

STAIN FILMS ON SILICON

R. J. ARCHER

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

Abstract—A surface reaction resulting in film growth occurs on silicon immersed in 10^{-1}–10^{-3} molar solutions of nitric acid, sodium nitrite or nitrogen dioxide in hydrofluoric acid. The films also form as the result of a standard etching procedure. Three types of films are obtained depending on the resistivity and type of the subphase silicon. On the basis of silicon content, refractive index and other properties, it is tentatively concluded that one type of film is elemental silicon. The other types may be silicon hydrides. The rate of film growth is proportional to the reciprocal of film thickness and to the 3/2 power of the initial concentration of nitric acid or sodium nitrite. The rate of growth is nearly independent of subphase resistivity except for *p*-type silicon in the resistivity range from 0·1 Ω-cm to 0·001 Ω-cm which encompasses a maximum in a plot of rate versus resistivity. The elemental silicon type films grow on subphases in this resistivity range while all other *p*-type material and all *n*-type subphases grow the hydride type films.

1. INTRODUCTION

THE occasional formation of colored stain films on silicon following etching in HF–HNO_3 solutions is commonly observed. FULLER and DITZENBERGER[1] established that such films can be consistently grown in dilute solutions of nitric acid in concentrated hydrofluoric acid; TURNER[2] and UHLIR[3] report similar film growth during the anodization of silicon in hydrofluoric acid, and the film growth on germanium in HF–H_2O_2 solutions observed by ELLIS[4] is probably a related phenomenon. This paper reports a study of the properties of stain films on silicon and of the staining process.

Before presenting the results, an instance of the occurrence of these films, of significance to those concerned with the preparation or study of silicon surfaces, will be described. Ellipsometer[5] measurements reveal that films 100–200 Å thick consistently form on silicon surfaces in HF–HNO_3 etching solutions when the dissolution action of the etch is quenched by dilution with water. Fig. 1 plots data showing this effect. Curve (1) is the result of a cleaning and etching procedure, developed to give surfaces with minimal film thicknesses, the final step of which requires quenching the HNO_3–HF etch by adding a large excess of nitric acid and then rinsing with water. If, instead, the etching action is stopped by diluting the etch with copious amounts of running distilled water, curve (2) results. The thicknesses of the films following water quenching are an order of magnitude larger than for nitric acid quenching, and

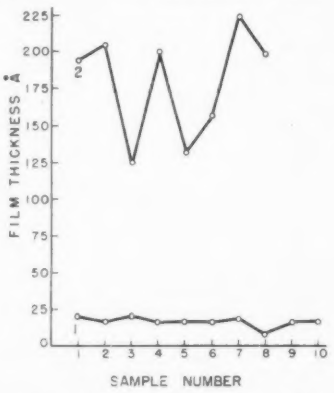


FIG. 1. Films on silicon surfaces following etching quenched by HNO_3, curve (1) and by water, curve (2). 1 Ω-cm *p*-type.

their chemical properties establish that they are the same substance as the thicker interference color range stain films.

The following sections summarize the physical and chemical properties of the stain films and of the staining process. The results include measurements of the composition, refractive index, density

and electron diffraction patterns of the films and determination of the kinetics of the staining reaction and the dependence of all of these properties on the composition and concentration of the staining solution and on the resistivity and type of the subphase silicon.

2. PROPERTIES OF STAIN FILMS

The appearance of the films depends on the resistivity and type of the subphase silicon. The films may be classified into three types on this basis. Type I films exhibit normal interference colors. They form on *n*-type material of any resistivity and on all *p*-type material except for that in the resistivity range from 0·25 Ω-cm to 0·001 Ω-cm. Type II films show no interference colors beyond a certain thickness but are black. These films grow only on *p*-type silicon in a narrow resistivity range including 0·25 Ω-cm and 0·1 Ω-cm. For these two materials, respectively, the films are black for thicknesses greater than those corresponding to the second order blue and first order brown interference colors. Type III films grow on *p*-type silicon in the approximate resistivity range from 0·1 Ω-cm to 0·001 Ω-cm. For subphases near the upper extreme of this range the films are invisible. They exhibit neither interference nor intrinsic colors at any thickness. As the subphase resistivity decreases in the type III range, very pale interference colors are observed which increase in intensity with decreasing resistivity. For subphases lower in resistivity than 0·001 Ω-cm, normal type I colors are observed.

The silicon content of the stain films was determined by two different methods. Firstly, the proportion of silicon in the film was measured by dissolving a weighed quantity of film in a dilute alkaline solution and determining gravimetrically the silica content of the solution by the standard analytical procedure. The results for films of types I and III are shown in the third column of Table 1. The indicated error of ±3 per cent was determined by applying the analysis to solutions of known silicon content.

The second method for obtaining the silicon content of the films was by measuring the quantity of oxygen taken up by the films at high temperatures in an oxygen ambient. When a silicon specimen uniformly covered with a stain film is heated in oxygen at 900–1000°C there is an initial rapid uptake of oxygen followed by a much slower subsequent uptake. The fast and slow uptakes result, respectively, from oxidation of the film and of the subphase. After the fast initial oxidation, the film has the chemical properties of silica and readily dissolves in hydrofluoric acid. By determining the masses of the specimen before oxidation, after oxidation and after removing the oxidized film in hydrofluoric acid, the ratio of the mass of the initial stain film to that of the oxidized film may be obtained. On the assumption that the oxidized film is pure silica, this ratio yields the proportion of silicon in the initial stain film. The results of such measurements for all three types of films are given in column two of Table 1. The indicated errors are the average deviations of several determinations.

Table 1. *Weight per cent silicon content of films*

Film type	Film oxidation (wt.%)	Silicon analysis (wt.%)
I	88±8	91±3
II	97±1	
III	102±5	100±3

Within experimental error, the two analytical methods agree. It is concluded that type III films may be elemental silicon whereas type II and more certainly type I films are silicon rich compounds or mixtures. At the same time as the present work, but independently, TURNER[6] found that a stain film grown anodically in hydrofluoric acid was about 96 per cent silicon.

The chemical properties of the films are similar to those of bulk silicon except that the films are more reactive. The films dissolve in aqueous alkaline solutions of pH greater than 9 at room temperature whereas bulk silicon is only attacked by hydroxide solutions at temperatures above 60–70°C. They are oxidized at high temperatures in oxygen and in warm nitric acid much more rapidly than bulk silicon. There are considerable differences in the rates of these reactions between the different types of film in the order

$$\text{I} > \text{III} > \text{II}$$

The films are insoluble in HF but are rapidly dissolved by silicon etches.

The refractive indices of the film were measured by the method of ABELES[7]. The method requires the determination of the angle of incidence at which monochromatic light polarized in the plane of incidence is reflected at the same intensity from the clean subphase and from the film covered surface. This angle is the Brewster angle for the film substance which is, of course, a measure of its refractive index. Experimental results are plotted in Fig. 2 as film refractive index at 5461 Å versus

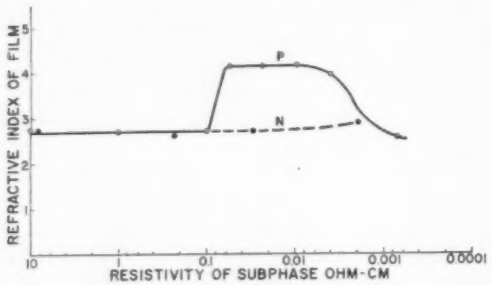


FIG. 2. Refractive index of film versus subphase resistivity.

subphase resistivity for both *n*- and *p*-type silicon. These optical properties account for the appearance differences between films of types I and III. The occurrence of type I films corresponds to the entire resistivity range giving the nearly constant refractive index of about 2·7.* Invisible and pale-colored type III films correspond to the broad maximum of the *p*-type curve at a refractive index value of about 4·1. The refractive index of single-crystalline silicon is 4·13.[5] Thus, the magnitudes of the optical discontinuities at the film–subphase interface explain the difference in appearance between type I and type III films, and the agreement between the type III film refractive index and that of bulk silicon supports the view that these films are elemental silicon.

Knowing the refractive index, film thickness can be determined from interference color using Rollet's table of optical thickness versus color.[8] Thickness values thus obtained combined with measurements of surface density (g/cm^2) give the densities of the various films. For type I films the density is 1·5 g/cm^3, for type III films (0·004 Ω-cm *p*-type subphase) 2·0 g/cm^3.

The results of electron diffraction measurements on film covered surfaces are ambiguous. Two measurements on type III films gave silicon single-crystalline arrays one of which also showed one sharp ring consistent with the pattern from poly-crystalline silicon. A type II film also gave the single-crystalline silicon pattern. One type I film gave a broad diffuse ring indicating an amorphous layer. Another type I film gave two interpenetrating single-crystalline arrays, one of which matches silicon; the other is weak and unidentified. One point of ambiguity is whether the single-crystalline arrays indicate the structure of the films or of the subphase. If the former, film growth is epitaxial.

3. PROPERTIES OF THE STAINING PROCESS

Films of uniform coverage and exhibiting interference colors through as many as 7 orders grow on freshly etched silicon immersed in 10^{-1}–10^{-3} molar solutions of nitric acid in concentrated hydrofluoric acid. The rate of growth increases with concentration and depends on subphase resistivity and type. Rates of hundreds of Å/sec down to fractions of Å/sec are observed. Film growth is also obtained in solutions of sodium nitrite in hydrofluoric acid and of nitrogen dioxide in hydrofluoric acid.

Some samples of hydrofluoric acid do not produce staining solutions even for nitric acid concentrations as high as 0·02 molar, and *any* sample of hydrofluoric acid is made inactive by exposure

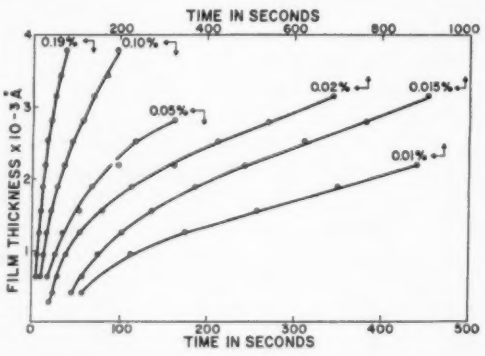


FIG. 3. Film growth rates on 0·004 Ω-cm *p*-type silicon as a function of nitric acid concentration in vol. %. Note two time scales.

* The point for 0·1 Ω-cm *p*-type was obtained for a film showing the first order brown color and may be considered of type I rather than type II.

VOL. 14 1960

to a large area of powdered silicon. It is assumed, that some impurity in the hydrofluoric acid is necessary to activate the staining reaction and that powdered silicon "getters" this impurity. Active staining solutions are very slightly brownish or blue-gray in color depending on concentration. Inactive solutions are colorless. "Gettering" has no effect on the staining action of sodium nitrite or nitrogen dioxide solutions.

Vigorous agitation or stirring stops film growth in nitric acid solutions but increases the growth rate in sodium nitrite solutions and is without effect on nitrogen dioxide solutions. The addition of urea to the staining solutions inhibits film growth. Intense illumination increases the growth rate on *n*-type silicon and decreases that on *p*-type. Normal room light does not affect the rate. Open beakers of the staining solutions decrease in strength with time.

The rate of film formation in room light at $25{\cdot}0 \pm 0{\cdot}2°C$ and the dependence on nitric acid concentration, expressed in volume per cent of concentrated nitric acid (70·6 per cent) in concentrated hydrofluoric acid (48 per cent), is shown in Fig. 3. The silicon is 0·004 Ω-cm *p*-type which gives films more or less intermediate between types I and III with pale interference colors through 7 orders. Thickness is obtained from color using the refractive index 3·96. The data obey a parabolic rate law reasonably well as shown by the linearity of the thickness-square plot of Fig. 4. The dependence of the parabolic rate constant on nitric acid concentration is obtained from the log–log plot of the slopes of Fig. 4 against

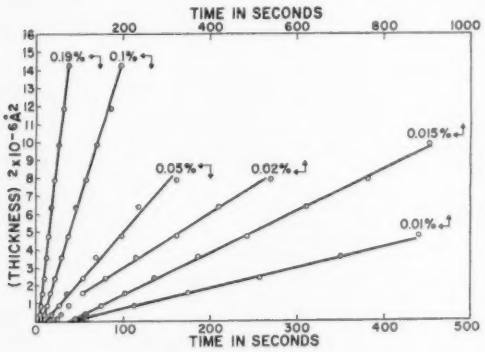


FIG. 4. Film growth rates, thickness square versus time, as function of nitric acid concentration. Note two time scales. 25°C.

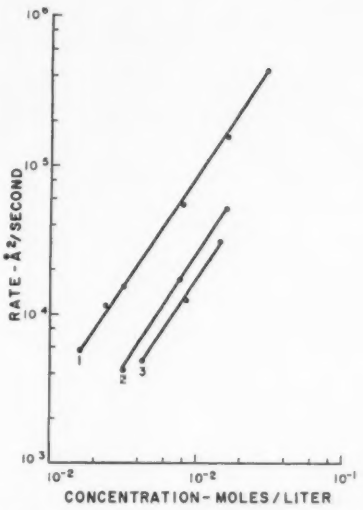


FIG. 5. Thickness square versus time slopes as a function of concentration. (1) 0·004 Ω-cm *p*-type in HNO_3 solutions. (2) 10 Ω-cm *p*-type in HNO_3 solutions. (3) 0·004 Ω-cm *p*-type in $NaNO_2$ solutions.

concentration, curve (1), Fig. 5. The slope of the curve is $1{\cdot}48 \pm 0{\cdot}02$, and all of the data of Fig. 3 may be expressed by

$$\frac{dL}{dt} = \frac{3{\cdot}73 \times 10^8 [HNO_3]^{1{\cdot}48}}{L} \qquad (1$$

where L is thickness in Å, time is in sec, concentration* in moles/l. and the boundary condition $t = t_0$ at $L = 0$ obtains. Time t_0 is the induction period seen in Fig. 4.

Solutions of sodium nitrite in hydrofluoric acid give data with the same characteristics as those of Fig. 3 and 4. The corresponding log–log plot of parabolic slope versus concentration for such data gives curve (3) of Fig. 5. The data obey relation 1 when the coefficient of that expression is multiplied by $(1/3)^{1{\cdot}48}$, i.e. to produce staining solutions of equal strength the molar concentration of sodium nitrite must be 3 times greater than that of nitric acid.

Data for film growth (type I) on 10 Ω-cm *p*-type silicon in HF–HNO_3 solutions is shown in Fig. 6. The lower curves exhibit a curious step-like

* Here and elsewhere "concentration" is used loosely to designate the number of moles of solute initially added to a liter of solvent.

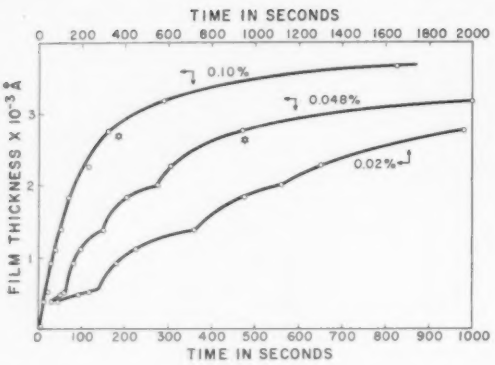


FIG. 6. Film growth rates on 10 Ω-cm *p*-type silicon as a function of nitric acid concentration in vol. %.

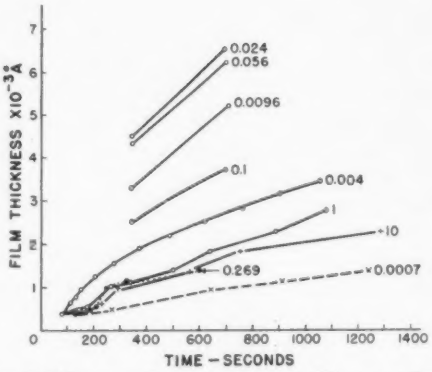


FIG. 7. Film growth rates on *p*-type silicon as a function of subphase resistivity, noted on curves in Ω-cm, in 0·015 vol. % HNO_3.

growth. This phenomenon is consistently observed for type I films in dilute staining solutions but not for type III films. Another unusual aspect of the data is the approach to a limiting thickness shown by the upper two curves. Beyond a certain thickness (the starred points) the slopes decrease much more rapidly than parabolically. This behavior which is attributed to dissolution of the film by the solution is always obtained with *p*-type subphases of resistivity greater than about 0·1 Ω-cm but has not been observed for any other case. The data of Fig. 6 give reasonably linear plots on a thickness-square versus time scale up to and including the starred points and for all points on the lower curve. The results of such a plot give curve (2) of Fig. 5. The data are consistent with equation (1) if the constant ($3{\cdot}73 \times 10^8$) of that equation is multiplied by 1/3.

The dependence of the growth kinetics on subphase resistivity for *p*-type* silicon in 0·015 vol.% nitric acid solutions is shown in Fig. 7. The points on the four upper curves for invisible type III and black type II films were obtained by weighing the silicon specimen before and after dissolving the films in 20 per cent sodium hydroxide solutions. Surface density is then converted to thickness by means of the film densities given above. For type II films the density is unknown and the thicknesses of the black films on the 0·1 Ω-cm specimen were obtained using 1·5 g/cm³ the value for type I films. Consequently these data may be in error by a constant factor.

* For all resistivities, the silicon is boron doped.

The dependence of growth rate on subphase resistivity is shown in Fig. 8 where thickness, from Fig. 7, after 500 sec in the staining solution is plotted against resistivity. Corresponding results for *n*-type subphases are also shown but the detailed data will not be given. The *n*-type thickness versus time data are very similar to the 1 and 10 Ω-cm *p*-type curves.

The resistivity range encompassing the maximum in the *p*-type curve of Fig. 8 corresponds to the range of type III film formation (except for the 0·1 Ω-cm point which is for a type II film). Very approximately the relative parabolic rate constants are obtained by squaring the thickness ordinate of Fig. 8. It is thus seen that at the maximum, type III films grow about 16 times faster than type I films. The relative differences in the growth rates on *n*-type and *p*-type materials for resistivities

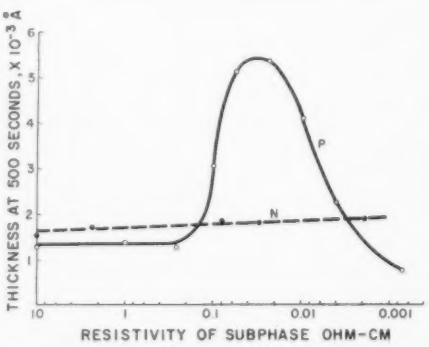


FIG. 8. Dependence of kinetics on subphase resistivity and type.

greater than 0·25 Ω-cm is real but is larger here than for similar experiments performed using more concentrated nitric acid staining solutions where the rate differences are almost within experimental error. The minimal rate for the 0·0007 Ω-cm *p*-type sample shown in Fig. 8 is observed for staining solutions of any concentration as is the small increase in rate with decreasing *n*-type subphase resistivity. The similarity of the characteristics of the refractive index curves, Fig. 2, and the rate curves of Fig. 8 is noteworthy.

4. DISCUSSION

There has been speculation that the stain film substance might be silicon monoxide, but the silicon content of the films precludes that possibility. A more likely hypothesis is that type I and perhaps type II films are silicon hydride and that type III films are elemental silicon or silicon rich mixtures of elemental silicon and silicon hydride. Solid, polymeric silicon hydrides with compositions varying from SiH to SiH_2 are described in the literature.[9, 10] Little work has been done on the properties of these substances, but the following pertinent observations are reported. Like stain films, the material dissolves in aqueous alkaline solutions. The compounds decompose to elemental silicon and hydrogen when heated to 1000°C. Consistent with this hydride property, type I films heated to 1000°C in vacuum disappear, that is, the subphase surface no longer shows an interference color. Analysis by mass spectrometry of the evolved gases from a type I film following such a heating gave about equal quantities of H_2 and H_2O. The total quantity of hydrogen evolved, both as H_2 and as H_2O, would indicate a formula $SiH_{1/2}$ for the particular film mass heated. However, this one run is not believed to be quantitatively valid because of the possibility that the film was partially oxidized (hence the evolved water) and that the film substance was not completely decomposed. Qualitatively, though, the result supports the hydride hypothesis. No H_2 or H_2O (beyond normal background) was obtained when a silicon specimen without a film was heated. Finally, no fluorine, nitrogen or oxygen was evolved from the film. Further measurements are planned to establish quantitatively the hydrogen content of the films. Within experimental error the proportion of silicon in type I films is consistent with the formula SiH_2. Finally, the formation of silicon hydride under staining conditions would not be a unique phenomenon since the formation of hydride films on a number of metals during dissolution in hydrofluoric acid has been observed.[12]

The facts that type III films have refractive indices very close to that of elemental silicon, that their silicon content is 100 per cent ± 3 per cent and that they are considerably less reactive than type I films, e.g. they dissolve much more slowly in alkaline solution, leads to the view that they are predominantly elemental silicon.

Consistent with UHLIR's[3] measurements on the anodic formation of stain films on silicon, the oxidation step in the film growth corrosion process is assumed to be

$$Si \rightarrow Si^{++} + 2e \qquad (2)$$

The resultant two-plus silicon compound could be SiH_2 for type I films, or for films which are elemental silicon it can be supposed that an intermediate two-plus compound forms and spontaneously decomposes to elemental silicon and some oxidation product. Most compounds which contain silicon in the two-plus oxidation state are thermodynamically unstable at room temperature.

In all of the staining solutions, the species NO_3^-, HNO_2, NO_2 and NO coexist. The cathodic or reduction step of a corrosion reaction in this system is, according to VETTER[11],

$$e + NO_2 \rightarrow NO_2^- \qquad (3)$$

and the rate limiting step is normally the formation of NO_2 by the reaction

$$HNO_3 + HNO_2 \rightleftarrows 2NO_2 + H_2O \qquad (4)$$

This mechanism explains the fact that nitric acid solutions made with "gettered" hydrofluoric acid do not stain, since the reaction cannot occur until it is initiated by at least a trace of NO_2 or HNO_2, and this trace comes from the reduction of nitric acid by impurities in the hydrofluoric acid which are removed by "gettering". The reaction is autocatalytic since once started, it leads to a build up in the concentration of HNO_2 at the interface. Agitation, in the case of nitric acid solutions, inhibits film growth by dispersing the large interfacial concentration of HNO_2. The opposite or nil

agitation effects in NO_2 and $NaNO_2$ solutions must be related to the fact that these solutions contain relatively large concentrations of HNO_2. Urea inhibits film growth by decomposing HNO_2 to N_2.

The 3:1 molar ratio between sodium nitrate and nitric acid staining solutions of equal strength may be explained by the following decomposition of the nitrite

$$3HNO_2 \rightarrow HNO_3 + 2NO + H_2O \qquad (5)$$

The parabolic kinetics of the staining reaction suggests that the rate limiting step is the diffusion of a reactant through the growing film, but the 3/2 power dependence on the initial quantity of nitric acid or of sodium nitrite used to make the solutions has no obvious interpretation.

The abrupt change in the properties of the film and in their rates of growth at the onset of the maxima of Figs. 2 and 8, may reflect the electrochemical nature of the corrosion process that causes film growth. Although a mechanism that explains the maxima has not been worked out, the following speculation may be pertinent. At those parts of the subphase surface which are cathodically polarized, a space charge barrier exists in the silicon which limits cathodic current, and if this current were rate limiting for film growth an abrupt increase in rate would be expected when the breakdown voltage of the barrier was smaller than the voltaic e.m.f. of the corrosion reaction. The breakdown voltage depends on resistivity and for material in the resistivity range of the maxima is about 1–2 volts.[13] This value is of the same order of magnitude as the e.m.f. of the reaction. The indicated mechanism could thus account for the increase in growth rate, but is apparently inconsistent with the subsequent decrease in rate at lower resistivities and with the fact that the kinetics indicate a rate limiting step involving diffusion through the growing film.

Clearly, the formulation of a detailed mechanism consistent with all of the properties of this unusual reaction and of the resulting films will require further experimental effort.

The author wishes to thank R. D. Heidenreich and Mrs. M. H. Read for the electron diffraction measurements, J. P. Wright for the silica analysis and E. E. Francois for the mass spectrometry determination.

REFERENCES

1. Fuller C. S. and Ditzenberger J. A., *J. App. Phys.* **27,** 550 (1957).
2. Turner D. R., *Trans. Electrochem. Soc.* **105,** 402 (1958).
3. Uhlir A., *Bell Syst. Tech. J.* **35,** 333 (1956).
4. Ellis S. G., *J. Appl. Phys.* **28,** 1262 (1957).
5. Archer R. J., *Trans. Electrochem. Soc.* **104,** 619 (1957).
6. Turner D. R., private communication.
7. Abeles F., *J. Phys. Radium* **11,** 310 (1950).
8. Constable F. H., *Proc. Roy. Soc.* A**115,** 570 (1927).
9. Stock A., *Hydrides of Boron and Silicon.* Cornell University Press, Ithaca, N.Y. (1933).
10. Schwarz R. and Heinrich F., *Z. anorg. allgem. Chem.* **221,** 277 (1934).
11. Grahame D. C., *Ann. Rev. Phys. Chem.* **6,** (1955).
12. James W. J. and Straumanis, *Trans. Electrochem. Soc.* **106,** 631 (1959).
13. Ansley W. G., unpublished data on breakdown voltage of silicon *p–n* junctions as a function of resistivity.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 111–116. Printed in Great Britain.

GAS ADSORPTION AND SURFACE CHARGE DENSITY OF GERMANIUM SURFACES

M. J. SPARNAAIJ

Philips Research Laboratories, Eindhoven, Netherlands

VOL. 14 1960

Abstract—The influence upon adsorption properties is studied of the presence of discrete charge units at the surface of an adsorbent assuming a model given by ESIN and SHIKOV[1] and by GRAHAME[2]. This influence is certainly small when the surface charge density is lower than 10^{13} electronic charges/cm^2 and it is also small when this density is higher than 10^{15} electronic charges/cm^2. The influence may be appreciable when the surface charge density is in between these two values. A possible application may be found in the interpretation of adsorption data of argon at an oxidized germanium surface, Such a surface then would carry 10^{14}–10^{15} electronic charges per cm^2 which are compensated by positive charges at a distance of some angstrom units in the direction of the bulk. Data concerning the chemisorption of oxygen at a non-oxidized germanium surface are communicated.

1. INTRODUCTION

PHYSICAL adsorption of noble gases on solid surfaces is usually explained by the London–Van der Waals attraction between the gas atoms and the atoms of the solid material. However, when electrostatic fields are present near the surface, another type of attraction may play a role, the attraction energy V being given by:

$$V = -\tfrac{1}{2}\alpha F^2 \qquad (1)$$

where α is the polarizability of the gas atom and F the field strength. It is the purpose of this paper to show, that in some cases this polarization energy V is of the same order of magnitude as the calculated London–Van der Waals energy and will lead to a contribution to the adsorption energy which may not be neglected.

Consider for this purpose a layer of discrete charges at the surface of an adsorbent, these charges being compensated by a space charge layer in the bulk at some distance from the surface. The layer of discrete surface charges in combination with the compensating space charge, will give rise to local electrostatic fields near the surface and it is only in the case of a continuous charge distribution both of the space charge and of the surface charges that these electrostatic fields exactly balance each other at any position in a plane parallel to the surface (Fig. 1, arrow). Therefore consideration of discrete charges is an essential feature of the treatment to follow. For convenience only the surface charges will be considered as discrete entities. The compensating space charge which is at a larger distance from the surface will be considered as smeared out. This seems justified since a layer of discrete charges

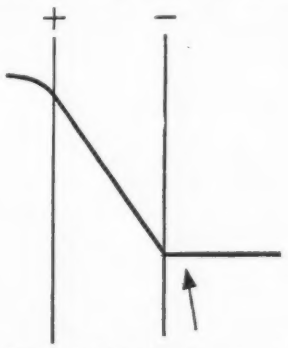


FIG. 1. Two layers of charge, both smeared out. A positive space charge is indicated. No field present at the position where the arrow is drawn.

and a layer with a continuous charge distribution, both having the same surface charge density σ, will both produce a field with a field strength equal to $2\pi\sigma$, if the distance is sufficiently large.

The discrete surface charges give rise to considerable variations of the field strength in a plane

parallel to the surface. Thus in a small area close to a surface charge unit (we consider the surface charges as negative and the compensating space charge as positive) the field strength has a high negative value and in the large remaining area the field strength is positive because there the field due to the space charge is incompletely screened off by the surface charges. The surface charges, together with cracks, holes and so on, will be considered as active spots. These active spots attract gas atoms in an early stage of the gas adsorption and will give rise to relatively high values of the adsorption energies at a low degree of coverage θ, say $\theta < 0{\cdot}1$. The field elsewhere may be considered as representative for the surface. Therefore we will be mainly occupied in this paper with estimates of this field. Reasonable estimates can be carried if an appropriate model is used. Such a model is available and is considered in detail by ESIN and SHIKOV[1] and by GRAHAME[2]. Their results are illustrated qualitatively in Fig. 2. A high surface charge density

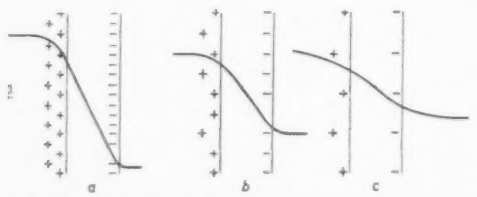


FIG. 2. Potential (vertical) versus distance (horizontal). A positive space charge is indicated in the figures.

($\approx 10^{16}$ electronic charges/cm^2) will only lead to a high field value close to the surface (Fig. 2a). The minimum distance between a gas atom and the surface may then be so large, that an appreciable value of V can not be attained. At the other extreme, a low surface charge density ($\approx 10^{13}$ electronic charges/cm^2) will lead to a field strength which extends to a larger distance, but whose value is small everywhere. It is only in the middle range, that one can expect the charges to play an important role. Here a sufficiently large value of the field strength ($\sim 10^6$ V/cm) may be found at a distance of say $2\frac{1}{2}$ Å, giving rise to a relatively large value of the polarization energy V.

Curves of the polarization energy versus distance will be compared with a "reference curve" calculated from the London–Van der Waals interaction energy. This interaction energy not only depends on such atomic characteristics as polarizability and ionization energy but also on the flatness of the surface. A rough surface leads to a higher value of the interaction energy than a smooth surface. Therefore in actual physical situations the London –Van der Waals attraction may be larger than expressed by our "reference curve". Although the comparison necessarily has a qualitative character, interesting details concerning the electric state of the surface can be revealed from it.

It should be emphasized that in the calculations to follow the field strength is considered to have a constant value in time. This is essentially different from calculations of Van der Waals forces where, in using equation (1), the time average value of F^2 is used, the time average value of F being zero. Now actual situations may be intermediate between these two extremes, i.e. it may be that the electric charges as depicted in Fig. 2 vary their positions in time such that the time average value of F is relatively small whereas the time average value of F^2 is considerable. Such a situation may occur, if the surface layer of an adsorbent consists of polar molecules. Although at present detailed calculations for this situation are not available, their results are likely to be qualitatively the same as those given in this paper.

2. CALCULATION

2.1. *The polarization energy*

Two contributions for the calculation of the electric field outside the surface can be considered: the contribution due to the presence of the charges on the surface and the contribution from the space charge compensating the surface charge (see Fig. 3). The field due to discrete surface charges can be calculated from a model proposed by ESIN and SHIKOV[1] and by GRAHAME[2]. This model consisted of circles with radius r_1 and a charge unit at the centres, the circles being arranged in a two-dimensional hexagonal pattern. A sheet of such circles was then considered. The sheet was made circular as far as possible and the charge at the centre of the sheet was assumed to be removed, the other charges retaining their position. The potential, ψ_0, say at the position of the missing charge, was determined by adding the coulomb-potentials of the charges in the sheet. Then the potential at a distance R of the sheet at the axis of the missing charge was computed in the same way. This potential be denoted as ψ_r. Whereas ψ_0 and

ψ_r separately do not converge to some limit as the area of the sheet is extended by taking more circles, the difference $\psi_0-\psi_r$ does and appears to be:

$$\psi_0-\psi_r = 0{\cdot}8\times 2\pi\sigma\{(r_1^2+R^2)^{1/2}-r_1\} \qquad (2)$$

where σ is the surface charge density, the surface charges being negative in sign. One has $\pi r_1^2\sigma = e$.

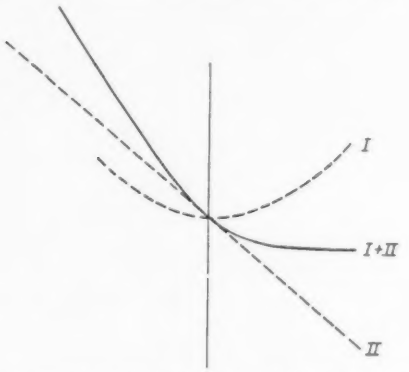


FIG. 3. Potential contributions near a flat plate with discrete charge units. The dielectric constant has the same value everywhere in the figure.

The expression holds good, if the surface area of the sheet is more than $\pi(10r_1)^2$ and if $0{\cdot}02<R/r_1<0{\cdot}5$. The numerical factor 0·8 becomes somewhat higher if $R/r_1>0{\cdot}5$. It tends to unity at still larger values. Then the potential difference approaches the usual expression $2\pi\sigma R$. According to Laplace's equation $\Delta\psi = 0$ and owing to the fact that the potential is periodical along the surface, the expression $2\pi\sigma R$ will be approached exponentially.

The field arising from $\psi_0-\psi_r$ is given by

$$-\frac{\partial(\psi_0-\psi_r)}{\partial R} = -2\pi\sigma\times\frac{0{\cdot}8R}{(r_1^2+R^2)^{1/2}} \qquad (3)$$

For the calculation of the field contribution due to the compensating space charge, it will be assumed that the distance between the individual charges and the position which is under consideration outside the surface, is large. Then this contribution is equal to two and the whole field becomes:

$$F = 2\pi\sigma\left(1-\frac{0{\cdot}8R}{(r_1^2+R^2)^{1/2}}\right) \qquad (4)$$

Image effects may interfere with the field given by equation (4). Consider two adjoining dielectrics, 1 and 2, with dielectric constants ϵ_1 and ϵ_1, respectively, and a charge e in dielectric 1. The field at a point in this dielectric can then be calculated as if only this dielectric were present, containing now an additional charge e' at the mirrored position. It follows from electrostatic theory that

$$e' = -e\frac{\epsilon_2-\epsilon_1}{\epsilon_2+3\epsilon_1} \qquad (5)$$

Thus, if in vacuum ($\epsilon_1 = 1$) a charge unit is considered very close to the surface of a dielectric with high ϵ_2 value, the field in vacuum will be small. If, now returning to the situation pictured in Fig. 2, the surface charges are located in dielectric 1 and the space charge in dielectric 2, the original assumption of a continuous charge distribution of the space charge is somewhat violated. Part of the space charge should be considered as concentrated in positions opposite the surface charge units and mirrored about the plane of separation of the two dielectrics. This leads to lower values of the field strength F. If $\epsilon_1 = 4$ and $\epsilon_2 = 16$ and if the distance between the surface charges and the plane of separation is 2 Å, this means that at a few angstrom units distance of the surface charges, the value of F should be lower by some 10 per cent compared with equation 4. Since there are two planes of reflection instead of one, a multiple reflection treatment should be followed according to ERSHLER[3] and to GRAHAME. Such a treatment will only lead to a small change of the calculated potential-distance curve in our case.

More important corrections are likely to arise from modifications of the model used by ESIN and SHIKOV and by GRAHAME[2]. It may be seen, that an irregular distribution of the surface charges leads to higher values of the field strength than a regular distribution because in the latter case the situation more resembles that of a continuous distribution than in the former. Now apart from the removal of the (important) central charge, the model used by ESIN and SHIKOV and by GRAHAME consisted of a regular array of charges. Therefore other models may provide higher representative values of the field strength.

2.2. *The London–Van der Waals energy*

POLANYI and LONDON[4] derived the following equation for the interaction energy E between a

gas atom and the (flat) surface of a solid containing q atoms per cm^3:

$$E = -\frac{q\pi\lambda}{6R^3} \tag{6}$$

where λ is a constant. This equation is obtained by an integration procedure, starting with the interaction energy v of two atoms, a distance r apart[5, 6]:

$$v = -\frac{\lambda}{r^6} \tag{7}$$

According to LONDON's harmonic oscillator-model,[5] the constant is given by:

$$\lambda = \tfrac{3}{4}\, h\sqrt{(V_1V_2)}\alpha_1\alpha_2 \tag{8}$$

where hV_1 and hV_2 are ionization energies of the atoms and where α_1 and α_2 are their polarizabilities.

Equation (6) can be improved by using a summation procedure[7] instead of an integration, and by taking into account, that the additivity of the interatomic interactions, underlying this summation or integration is not entirely justified.[8] Then the values of E become about twice as high. However, a much more important correction arises from the fact that a solid surface is never atomically smooth as assumed here. This is a difficulty inherent in all adsorption treatments and it cannot be accounted for in detail.

2.3. *Comparison*

In Fig. 4 four curves are shown; three curves are made according to equations (1) and (4) using three different values of σ and one curve is made according to equation (6) where $\lambda = 10^{-59}$ erg cm^6 and $q = 4\times10^{22}$ cm^{-3} were chosen. These are representative values for a number of cases.[9] It appears from a comparison between the curves, that the polarization energy can become of the same order of magnitude as the London–Van der Waals energy. It should be kept in mind that especially the London–Van der Waals energy, but also the polarization energy may have higher values than those calculated here from simplified models.

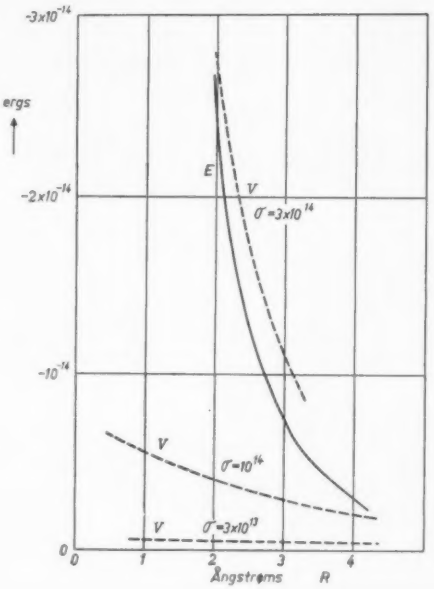


FIG. 4. Drawn curve according to equation (6). Dotted curves according to equations (1) and (4). The surface charge σ is expressed in number of electronic charges per cm^2. For $\sigma < 3\times10^{15}$ the interaction energy V will not extend beyond 3 Å. The polarizability of the argon atom was inserted in V and taken to be 1·65 Å^3. The constant $q\pi\lambda$ in E was taken to be 12×10^{-37} erg cm^3.

3. THE GERMANIUM SURFACE AND THE OXIDIZED GERMANIUM SURFACE

An application of the given calculation may be found in the interpretation of argon gas adsorption data, published elsewhere,[10] obtained at 77°K and 90°K with the germanium surface and the oxidized germanium surface as the adsorbents. In these experiments the germanium surface was obtained by crushing a crystal in air and, after bringing the obtained powder in a vacuum system to be used for volumetric adsorption measurements, heating at 600–800°C for 50 hr at least, the vacuum being less than 10^{-6} mm, or in a refined version of the apparatus (double walled quartz, Alpert–Bayard gauges) less than 10^{-9} mm. Dissolved gas and a surface oxide layer were removed by this treatment. After cooling the surface properties of the powder were much the same as those reported by other authors[11–15] who

VOL. 14 1960

prepared their surface by crushing a crystal in high vacuum or by other methods in absence of oxygen gas. A monolayer of oxygen is easily formed at 77°K or at room-temperature provided the oxygen gas pressure is sufficiently low ($<10^{-5}$ mm.) A second layer can loosely be bound at low temperatures and becomes firmly attached to the surface at a higher (i.e. room-) temperature. Oxygen gas pressures which are higher than say 1 mm result in a non-stoichiometric additional adsorption at room temperature and, at 250–300°C, in a third oxygen layer. This third layer is possibly the same as that reported by SUN[16] and may be due to a first step of oxygen diffusion into the bulk.[17] The oxygen can be removed probably in the form of volatile GeO.[18] This removal takes place at at least 575–600°C. The GeO settles on cold parts of the apparatus. According to X-ray investigations these (brownish) settlements were partly amorphous and consisted partly of crystalline Ge and GeO_2. Hardly any desorption takes place at room-temperature. This is in agreement with observations of LAW[19], ROSENBERG[13], WOLSKY[20] and SANDLER[15] all using different techniques. However, desorption at room temperature was found by FARNSWORTH[12] and DILLON[21]. Apart from this discrepancy which is not understood,[17] it is possible that all the authors mentioned have worked with the same type of surface.

Less chemisorption of oxygen gas than reported here may take place after a heat treatment at 900–950°C in bad vacuum or after etching the surface with oxidizing etchants.

Argon gas was chosen as the adsorbate instead of a heavier noble gas because its polarizability is relatively small and consequently the interaction between argon atoms and the surface could be studied without being obscured too much by an interaction according to equation (7) between the argon atoms themselves. Neon could not be chosen because the adsorption at 77°K is too low. The argon adsorption at 77°K and 90°K on a non-oxidized (i.e. regenerated at 600–800°C) surface was less than the adsorption on a surface containing two or three oxygen layers, the adsorption on the two latter, oxidized, surfaces being almost the same. This holds for a degree of coverage of argon atoms which is lower than unity. From the adsorption isotherms and thermodynamic analysis at a degree of coverage θ where $0{\cdot}1<\theta<0{\cdot}5$ it appeared that

$$\begin{aligned}\mu_1-\mu_2 &= 300 \text{ cal} \quad (77°\text{K})\\ &= 120 \text{ cal} \quad (90°\text{K})\end{aligned}$$

where $\mu_1-\mu_2$ is the difference of the chemical potentials of the adsorbed argon at the same degree of coverage on the non-oxidized (1) and the oxidized (2) surfaces. Furthermore, using adsorption models, which according to EVERETT[22] probably are oversimplifications, the difference of adsorption energies E_1-E_2 was estimated to be of the order of 1 kcal/mole, a reasonable estimate of E_1 giving 3 kcal/mole. It is this difference E_1-E_2 which may be attributed to a polarization energy as considered in this paper. In order to make sense, the polarization energy should be small at the non-oxidized surface, i.e. less than 10^{13} discrete charges/cm^2 should be present on the surface, and large at the oxidized surface, i.e. 10^{14}–10^{15} discrete charges/cm^2 should be present. This number compares favourably with the surface charge density as calculated by GREEN[23] from the shift of the contact potential after oxidizing a germanium surface. GREEN's interpretation was that this charge was smeared out over the surface.

This number may represent the number of active spots at the surface (see Introduction). An alternative interpretation can be found in the principles outlined in the last paragraph of the Introduction.

4. DISCUSSION

Although there is agreement with GREEN's conclusion concerning the surface charge density, alternative explanations for the different adsorptions may of course not be ruled out. Thus, the closest distance of approach of an argon atom to the surface may be smaller in the oxidized case than in the non-oxidized case. This is unlikely because the radius of an oxygen ion is larger than the radius of a germanium atom. Another possibility, which is more likely, is that the constant λ of equation (8) is larger for an oxidized surface than for a non-oxidized surface. Also, it is useful, according to equation (8), to compare the various polarizabilities and the ionization energies. Now the polarizability $\alpha_{O^{--}}$ of an O-ion is about 2 Å[3] and that of the Ge^{4+} is about 1 Å.[24] The value of α_{Ge} (the polarizability of the Ge-atom) is certainly higher. This also appears from the high refractive

index of germanium, which is about 4 at a wavelength λ_1 of the incident light of 0·5 μ and which is about 2 at $\lambda_1 = 0{\cdot}2\ \mu$.[25] Lorentz' equation then would give the result that $\alpha_{Ge} > \alpha_O$--. The constant λ (equation 8) depends also on ionization energies. Often the values for the atoms are inserted. Then $hV_{Ge} \approx 8$ eV and $hV_O \approx 13{\cdot}6$ eV. On the whole it is doubtful to predict that the constant for an oxidized surface is higher than that for a non-oxidized surface.

We finally return again to a possible influence of electrostatic fields. Their *origin* will not be discussed. It may merely be remarked that the discreteness-of-charge as the origin of electrostatic fields at the surface of an adsorbent has first been suggested by DE BOER[26] and later on by VERWEY[27] in a discussion of the surface properties of colloidal particles. Furthermore, HERMANN and WAGENER[28] and TEMKIN[29] considered an electronic gas close to the surface of metals as the origin of electrostatic fields.

As a general rule an electrostatic field will polarize adsorbed gas atoms. This polarization should lead to a shift of the contact potential. Such a shift has been found by MIGNOLET[30] for the case of noble gases adsorbed at metal surfaces. If discreteness-of-charge effects apply, a shift may also be expected in the case of the adsorption of (noble) gases at an oxidized germanium surface. However, the relation between this shift and the degree of coverage will be rather complicated since the field is not uniform along the surface and since mutual depolarization of the adsorbed gas atoms may become important at an increasing degree of coverage. Electrostatic fields will also influence the chemical potential- and the entropy-values of the adsorbed phase as has been pointed out by KEMBALL[31] and by DE BOER[32]. Although a clearcut interpretation of thermodynamic data is often difficult (EVERETT[22]), adsorption in therms of a variety of gases, in combination with measurements concerning the contact potentials may decide which mechanisms are responsible for the adsorption.

REFERENCES

1. ESIN O. A. and SHIKOV V. M., *Zh. fiz. khim.* **17,** 236 (1943).
2. GRAHAME D. C., *Z. Elektrochem.* **62,** 264 (1958).
3. ERSHLER B. V., *Zh. fiz. khim.* **20,** 679 (1946).
4. POLANYI M. and LONDON F., *Naturwissenschaften* **18,** 1099 (1930).
5. LONDON F., *Z. Phys. Chem.* B**11,** 222 (1930).
6. SLATER J. C. and KIRKWOOD J. G., *Phys. Rev.* **37,** 682 (1931).
7. DE BOES J. H. and CUSTERS J. F., *Z. Phys. Chem* B**25,** 225 (1934).
8. SPARNAAIJ M. J., *Physica, s' Grav.* **25,** 217 (1959).
9. MARGENAU H., *Rev. Mod. Phys.* **11,** 1 (1939).
10. SPARNAAIJ M. J. reported at *Brussels Conference on Semiconductors*, June 1958.
11. GREEN M., KAFALAS A. and ROBINSON P. H., *Semiconductor Surface Physics*, p. 349. University of Pennsylvania Press, Philadelphia (1957).
12. SCHLIER R. E. and FARNSWORTH H. E., *Semiconductor Surface Physics*, p. 3. University of Pennsylvania Press, Philadelphia (1957).
13. ROSENBERG A. J., ROBINSON P. H. and GATOS H. C., *J. Appl. Phys.* **29,** 771 (1958).
14. DELL R. M., *J. Phys. Chem.* **61,** 1584 (1957).
15. SANDLER Y. L. and GAZITH M., *J. Phys. Chem.* **63,** 1095 (1959).
16. SUN R., *Bull. Amer. Phys. Soc.* ser. 1, **1,** 144 (1956).
17. See also SCHLIER R. E. and FARNSWORTH H. E., *J. Chem. Phys.* **30,** 917 (1959) for the communication of preliminary results obtained by WOLSKY S. P.
18. LAW J. T. and MEIGS P. S., *Semiconductor Surface Physics*, p. 383. University of Pennsylvania Press, Philadelphia (1957).
19. LAW J. T. and GARRETT C. G. B., *J. Appl. Phys.* **27,** 656 (1956).
20. WOLSKY S. P., *J. Appl. Phys.* **29,** 1132 (1958).
21. DILLON J. A. and FARNSWORTH H. E., *J. Appl. Phys.* **28,** 174 (1957).
22. EVERETT D. H., *Proc. Chem. Soc., Lond.* p. 38, Feb. (1957).
23. GREEN M., reported at *Brussels Conference on Semiconductors*, June 1958, p. 362.
24. TESSMAN J. R., KAHN A. H. and SHOCKLEY W., *Phys. Rev.* **92,** 890 (1953).
25. Data compiled in GMELIN's *Handbuch der anorganischen Chemie*, Ge Ergänzungsband, Syst. no. 45 (1958).
26. DE BOER J. H. and VEENEMANS C. F., *Physica, s' Grav.* **1,** 960 (1934). Also DE BOER J. H., *Electron Emission and Adsorption Phenomena*, University Press, Cambridge (1935); *Proc. Symp. Keele*, 1956, p. 27.
27. VERWEY E. J. W., *Rec. trav. chim.* **53,** 933 (1934).
28. HERMANN and WAGENER, *Die Oxydkathode* Leipzig (1948).
29. TEMKIN M. I., *Symposium on Problems of Chemical Kinetics, Catalysis and Reactivity*, Moscow 1955. See also DE BOER J. H., *Chemisorption* in *Proc. Symp. Keele*, 1956, p. 27.
30. MIGNOLET J. C. P., *Disc. Faraday Soc.* **8,** 105 (1950); *J. Chem. Phys.* **21,** 1298 (1953). For data on clean germanium, see Ref. 22.
31. KEMBALL C., *Advanc. Catalys.* **2,** 233 (1950).
32. DE BOER J. H., *Advanc. Catalys.* **8,** 17 (1956).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 117–123. Printed in Great Britain.

CALORIMETRIC DETERMINATION OF THE HEAT OF ADSORPTION OF OXYGEN ON EVAPORATED FILMS OF GERMANIUM AND SILICON

D. BRENNAN, D. O. HAYWARD* and B. M. W. TRAPNELL†

Department of Inorganic and Physical Chemistry, University of Liverpool, England

VOL. 14 1960

Abstract—The heats of adsorption of oxygen on evaporated films of germanium and of silicon have been determined to an absolute accuracy of ± 5 per cent by means of a Beeck type calorimeter. The variation of the heat of adsorption with surface coverage has been studied, together with the simultaneous change in film resistance. The significance of the results is discussed.

INTRODUCTION

THE primary object of this investigation has been the determination of the heats of adsorption of oxygen on germanium and silicon; in addition, some subsidiary observations concerning the adsorptions have also been made. GREEN *et al.*[1] have already indicated the need for a reliable calorimetric investigation of these systems. For their work with germanium, these authors used crushed samples and the temperature rise following the adsorption was measured by means of a thermistor gauge embedded in the sample. This arrangement suffers from two major limitations. Firstly, temperature gradients within the calorimeter will be large and their presence will give rise to error both in the actual adsorption determination and, again, in the subsequent heat capacity determination. Secondly, difficulty arises from the appreciable time taken for the gas to diffuse into the tightly packed adsorbent. For the present determinations, the adsorptions were made on evaporated films deposited in a Beeck type calorimeter.[2] The advantages and limitations of this design of calorimeter are referred to in the Discussion.

* Present address: Department of Chemistry, Imperial College of Science and Technology, London, S.W.7., England.

† Present address: Denstone College, Uttoxeter Staffordshire, England.

EXPERIMENTAL

The calorimeter

The full details of the calorimeter design and operation are to be published elsewhere in connection with a separate study; the following is a brief account of the main features. The calorimeter consisted of a very thin walled glass tube wound with platinum wire which was made one arm of a Wheatstone bridge and which served as a resistance thermometer. The calorimeter tube was mounted vertically within a jacket through which a swift stream of water could be passed when a film was being laid inside the calorimeter tube, and which could be evacuated to prevent collapse during bake out and to reduce heat losses during a heat measurement. It was essential to good calorimetric technique to ensure that equal quantities of heat liberated anywhere within the calorimeter gave rise to equal changes in the resistance of the platinum thermometer. To this end, the calorimeter tube was prepared by the carefully controlled etching of a selected length of tubing and was of good longitudinal uniformity. The lateral uniformity, however, was not so satisfactory and its possible effect on the form of the results for silicon is referred to in the Discussion. The platinum thermometer was wound on a lathe to ensure accurate spacing of the turns. The part of the tube determining the heat capacity of the calorimeter was 17 cm long and 30 mm in bore and was positioned so as to ensure the absence of errors arising from

end effects. The films were confined to the central 10–11 cm section so that the platinum windings extended above and below the film. The overall sensitivity was approximately 1·7 m galvanometer scale deflection per degree. Scale deflections due to gas adsorptions varied from 4 to 20 cm/dose, the maximum deflection being reached in about 30 sec. The temperature drift rate was never more than 1 mm scale deflection/min and it was frequently much smaller than this. The cooling was accurately Newtonian in form, but was due almost entirely to radiation losses. The heat capacity of the calorimeter was $2{\cdot}84 \pm 0{\cdot}06$ cal deg^{-1}. It was obtained as the result of a total of 35 calibration measurements in which heat was developed electrically in a number of metal films deposited in the calorimeter and use was made of modifications of both the pulse method of BAGG and TOMPKINS[3] and of the equilibrium method of WHABA and KEMBALL[4].

Film resistance measurements

The films used for these measurements were deposited in a cylindrical vessel of ordinary glass tubing (3 cm bore) into the wall of which had been inserted two platinum ring electrodes at a separation of 9 cm; these electrodes consisted of 0·5 mm platinum wire spot-welded to tungsten leads. Measurement of its electrical resistance was made by making the film one arm of a Wheatstone bridge but, because the film resistances were in the region of 100 mΩ, the total resistance of the balance arm was only a few ohms, so care was taken to include the resistance of the connections when making an estimate of the absolute film resistance. Changes in film resistance were followed by observing the out of balance galvanometer deflection produced by an adsorption; this deflection varied linearly with change in resistance for changes of the magnitude occurring in these experiments.

Surface area determinations and general procedure

The pumping and gas dosage arrangements were similar to those which have been described in earlier publications.[5] Both the germanium and the silicon were evaporated from small pieces of the material lodged in tight coils made in 0·3 mm diameter tungsten hairpin filaments which had previously been thoroughly outgassed; each filament had from six to eight such coils. Due to the comparatively high resistance of the semiconductors, the coils were not shorted by their presence and temperatures adequate for the purpose of outgassing and evaporating the material could be achieved with moderate currents. Germanium was outgassed in the molten state before being placed in the calorimeter. The outgassing of the silicon was more difficult since silicon evaporates before melting; it also alloys with tungsten with consequent premature burn out of the filament if the temperature is not confined to narrow limits. To assist outgassing of the silicon, the pieces were raised for a short time to a sufficiently high temperature to cause them to disintegrate and become distributed as a thin layer over the surface of the filament. The germanium used was 10–15 Ω-cm *n*-type and the silicon was 9 Ω-cm *n*-type; we are grateful to Bell Telephone Laboratories, Inc., for the gift of both materials. The filaments were positioned as centrally as possible in the calorimeter tube in an effort to ensure a uniform deposit, but, in the case of silicon, the films obtained were somewhat uneven in spite of the precautions taken to ensure uniformity. The films were laid only after rigorous outgassing of the apparatus and it is estimated from the observed outgassing rates that contamination of a film would never exceed a few per cent of the available surface. The films were sintered at 40°C immediately after deposition.

Oxygen was obtained by thermal decomposition of potassium permanganate. The surface areas of the films were measured by the B.E.T. method using krypton adsorption at −195°C; the krypton was obtained in break-seal vessels from the British Oxygen Co. Ltd. and had a purity of 99–100 per cent, any impurity being xenon. The area measurements were made on films deposited in the calorimeter and used subsequently in adsorptions. A value of 3·0 mm was used for the pressure of krypton at saturation at −195°C and the relative pressures employed did not exceed 0·03. A value of 19·5 Å^2 was employed for the area of the adsorbed krypton atom.

RESULTS

The calorimetric measurements

The dependence of the heat of adsorption of oxygen on germanium and silicon is shown in Fig. 1. Each increment of oxygen gave rise in effect to

an integral heat over a small range of coverage and the points plotted correspond to the centre of each such increment. To bring the heat curves obtained in different experiments on the same system into accurate alignment with respect to coverage, each set of data was normalized to the *average* value of

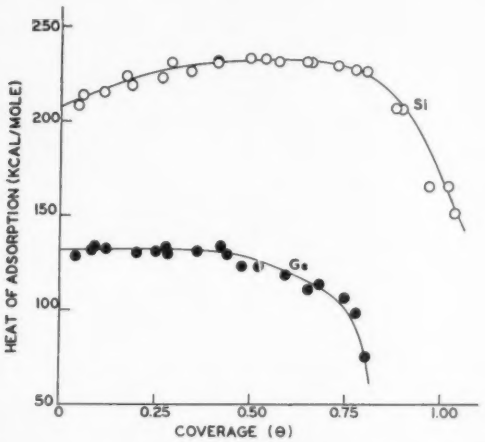


FIG. 1. The variation of the heat of adsorption of oxygen on germanium and silicon with surface coverage; the integral heats are 122 and 218 kcal mole^{-1}, respectively. The θ values refer to initial film areas. ● germanium, three determinations; ○ silicon, two determinations.

the volume of oxygen adsorbed per unit weight of film. Since the rate of oxygen uptake diminishes progressively with successive increments and there was no abrupt change in the rate at any particular stage in the adsorption, the limit of uptake used for normalization purposes in each case was taken as the coverage when the rate of adsorption had fallen to 6 μ ml (at s.t.p.)/min at a pressure of 20 μ; the average values used for the adsorptions were based on eight determinations in the case of germanium and six in the case of silicon. The θ values were calculated on the basis of 7·77 and $8{\cdot}36 \times 10^{14}$ adsorption sites/cm^2 for germanium and silicon respectively, θ being defined here as the number of oxygen *atoms* adsorbed per surface adsorbent atom. The number of surface adsorbent atoms per unit area for both germanium and silicon were based on a mean for the (111), (110) and (100) planes in which the first two planes were given a weight of two.[1] The mean maximum θ values corresponding to our criterion for the conclusion of the fast adsorption were 0·83 and 1·07 for germanium and silicon, respectively; the scatter on these values was not more than 2 per cent. A partially covered silicon film was left overnight, and the rate of uptake of oxygen and the heat of adsorption observed the next day corresponded precisely to the measurements made prior to the films standing; the same was true for the rate of adsorption on germanium, but no heat measurement was made in this case. This is good evidence that the surfaces had not suffered any significant contamination during this period which was considerably longer than normally taken for a determination.

Heat measurements with the present calorimeter were limited to fast adsorptions, and calorimetric measurements were discontinued, therefore, when the residual pressure of oxygen exceeded 10^{-6} mm after 3 min from the time of admission of the increment. The overall uncertainty, including both the irreproducibility arising in the calorimetric technique and that due to variation in film properties does not exceed 5 per cent and is probably rather less.

Reduction in film area due to oxygen adsorption

There is a reduction in film area of about 20 per cent for germanium and 30 per cent for silicon following the adsorption of oxygen. This effect is exemplified in Table 1 which contains a selection of typical data relating to the adsorptions. It is seen that while the the film areas per unit weight are variable, the θ values referred to the *initial* area are reproducible and these are the θ values used throughout the work. Also, the extent of area reduction due to the adsorption appears to be independent of the initial area.

Adsorption (d) on germanium was carried out in two stages. Oxygen was first admitted to the film at −196°C, when 0·010 ml was adsorbed; on warming the film to room temperature, a further 0·009 ml of oxygen was adsorbed.

The rate of the adsorption

A very rough estimate of the rate of adsorption could be obtained from the time taken by the calorimeter galvanometer to reach peak deflection. Some allowance was necessary for the response time of the system and this was estimated to be

Table 1. The adsorption of oxygen on germanium and silicon

Film weight (mg)		Volume of θ oxygen from adsorbed initial area (ml at s.t.p.)		Krypton monolayer volume before oxygen adsorption (ml at s.t.p.)	Krypton monolayer volume after oxygen adsorption (ml at s.t.p.)	Initial film area (m^2g^{-1})	Final area as percentage of initial area
Germanium							
(a)	8·3	0·074	0·80	0·123	0·098	78	80
(b)	109·5	0·482	0·82	0·781	0·555	37	71
(c)	45·6	0·315	0·83	0·500	0·380	57	76
*(d)	3·2	0·019	0·81	0·031	0·027	51	87
Silicon							
(a)	—	0·065	1·08	0·074	0·055	—	74
(b)	1·2	0·030	1·09	0·034	0·025	149	74
(c)	7·6	0·276	1·05	0·323	0·234	222	73

* See text.

15 sec from observations of adsorptions in which the apparatus and not the adsorption was the rate determining factor; in Fig. 2, t is the time to peak deflection less the response time. This quantity is seen to diminish from the very first dose in the case of germanium, but, in the case of silicon, the rate of adsorption remains immeasurably fast until about $\theta = 0{\cdot}5$, when it commences to decrease in the same characteristic way as for germanium.

Film resistance measurements

The resistance of films of germanium and silicon were determined after successive doses of oxygen until the films were saturated; the results are shown in Fig. 3. For germanium, the adsorption resulted first in an almost instantaneous change in resistance which was then followed by a slow change. Silicon did not show this effect; each adsorption resulted in a rapid re-adjustment of film resistance until the slow effects accompanying the presence of a residual pressure of oxygen in the system were reached.

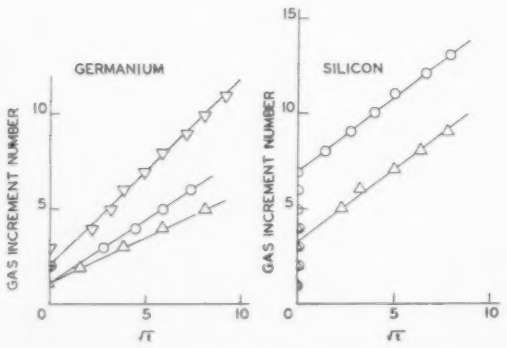


FIG. 2. Dependence on coverage of the rate of adsorption of oxygen on germanium and silicon; the differently styled points denote separate experiments.

DISCUSSION

Germanium

The maximum heat of adsorption of oxygen on germanium is 132 kcal $mole^{-1}$ which is very comparable to the heats of oxidation which are 129 and 122 kcal $mole^{-1}$, respectively, for GeO_2 and GeO[6]. A coincidence of the maximum heat of adsorption with the heat of oxidation has been observed by us for a number of the transition metals and will be commented on elsewhere. The initial heat of adsorption of oxygen on germanium observed by GREEN *et al.*[1] was in excess of 250 kcal $mole^{-1}$ and even allowing for the different nature of their adsorbent and ours, in that the distribution of crystal faces might not have been the same in the two cases, their value is undoubtedly too high for the reasons indicated in the Introduction. There is also a marked difference between the shape of these workers' heat of adsorption versus coverage curve which was sigmoidal, and that reported here. GREEN and LIBERMANN* have shown recently that

* private communication.

VOL. 14 1960

diffusion of oxygen into crushed germanium can be slow relative to the rate of adsorption and this will give rise to error both in the magnitude of their heat value, because the temperature of the thermistor will not be the same as that of the various parts of the sample, and additionally in the *shape* of the

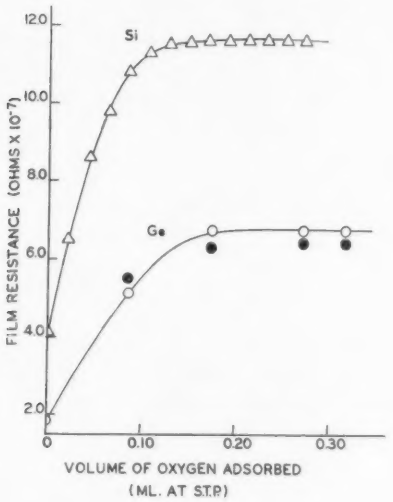


FIG. 3. The variation of film resistance with the extent of the adsorption of oxygen. ○, ● germanium; ○ denotes the final resistance and ● denotes the value to which the resistance first changed rapidly after the adsorption. △, silicon.

heat versus coverage curve. The *magnitude* of a heat of adsorption as determined with our calorimeter is independent of diffusion effects, so long, that is, as the adsorption is sufficiently rapid for the cooling correction to be relatively small and this is always so. On the other hand, the apparent variation of heat of adsorption with coverage is, of course, still a function of any diffusion effects. The fact that the curves relating heat of adsorption and film resistance with coverage both show a very definite change in trend at approximately $\theta = 0{\cdot}4$ suggests that the adsorption is a composite process. Against this must be set the observation that the rate of adsorption diminishes from the very beginning of the adsorption and is apparently determined by the same law throughout. However, the fact that the adsorption can be composite is clearly demonstrated by the observation that an adsorption carried out at $-196°C$ stops at $\theta = 0{\cdot}4$ and only at a higher temperature does the adsorption proceed further. The value $\theta = 0{\cdot}4$, which appears to have a special significance, may conceivably be related to the fact that 40 per cent of the total area is the contribution assigned to both the (111) and (110) surfaces. The idea of a composite adsorption has been used by HANDLER to account for changes in surface conductivity when oxygen is adsorbed on a *p*-type germanium surface. He postulated a relatively strong bond involving the dangling bonds of the (111) face and a relatively weak peroxide type bonding involving the (100) face. The present situation does not appear to be amenable to so simple a description. Evidently we are concerned with an *n*-type surface since adsorption of oxygen results in an increase in film resistance, but, in this connection, it is to be noted that the behaviour of germanium towards oxygen seems to be independent of whether the material is *n*- or *p*-type.[1] We have also to take into account the manner in which the heat of adsorption varies with coverage. This is determined by several factors which include the distribution of crystal faces within the film, the relative rates of adsorption on the various faces and the manner of their variation with coverage, and the total velocity of adsorption relative to the rate of diffusion into the film. Very little is known about any of these factors. A study of the adsorption of oxygen on the transition metals has shown that there are three limiting ways in which each gas increment may become distributed throughout the calorimeter. The gas may have sufficient time before it becomes adsorbed to pervade the whole of the adsorbent uniformly; or it may have sufficient time prior to adsorption to penetrate only part the way into the film, but to a uniform extent along the full length of the cylindrical sample, with the result that each increment becomes distributed in a thin layer which is located at a greater depth within the film with each successive dose. Finally, the gas may have a sufficiently high sticking coefficient to prevent its becoming uniformly distributed along the length of the calorimeter, but instead to cause it to condense in bands, each increment penetrating further down the calorimeter. Now, there is evidence[7, 8] to show that the adsorption of oxygen on nickel occurs in layers and for this system, the sticking

coefficient is approximately 5×10^{-3}[9] SCHLIER and FARNSWORTH[10] have reported a sticking coefficient of about 2×10^{-3} for the germanium (100) face and, since germanium films are more porous than those of nickel, a layering effect with germanium films seems probable. The sticking coefficient for the clean (111) germanium face is some 10 to 20 times smaller than for the (100) face. Hence, a possible explanation of the initial steady heat is that adsorption occurs preferentially on the (100) face to full coverage (or, possibly, on the (110) face also and then to approximately half coverage) so that the same integral heat would be observed for each successive layer. When the layering effect was complete, the remaining area would start to be covered and the heat would fall. The most recent electron diffraction evidence of SCHLIER and FARNSWORTH[10] is in favour of a one to one correspondence between adsorbed oxygen atoms and germanium lattice positions and any alternative explanation of the shape of the heat curve in terms of a direct or indirect interaction between such adsorbed atoms at a critical coverage will of necessity, in our present state of knowledge, be virtually entirely hypothetical.

In the absence of additional evidence, the most that it seems can be deduced with certainty from the present calorimetric measurements is that the heat of adsorption of oxygen on germanium at moderate coverages is comparable to the heat of oxidation and in attempts to describe the bonding on such surfaces it will not be necessary to seek processes which are exceptionally energetic.

The characteristic nature of the value of θ corresponding to the end of the fast adsorption, and the agreement of our value for this quantity (0·83) with that of GREEN *et al.*,[1] who found $\theta = 0{\cdot}9$ at the end of 6 sec is good evidence for the validity of our θ values as calculated on the basis of initial areas. The reduction in area following adsorption is probably entirely restricted to the area of the adsorption, the clean area being quite unaffected.

Silicon

Again there is fairly good agreement between the maximum heat of adsorption observed here (230 kcal mole^{-1}) and the heats of oxidation to SiO and SiO_2, both of which are approximately 210 kcal mole^{-1}, and again, therefore, the bonding in the surface layer is probably very similar to that found in the bulk oxide. The adsorption on silicon is much faster than on germanium and it is not until about $\theta = 0{\cdot}5$ that the time for an adsorption falls below the response time of the apparatus. This is in keeping with the higher sticking coefficients which have been reported for the oxygen–silicon system. Thus, EISINGER and LAW[11] find a value between 0·01 and 0·02 for an unoriented clean surface, while SCHLIER and FARNSWORTH report values of 0·02 for the clean quenched (100) surface and between 0·1 and 0·2 for the (111) surface. These coefficients are sufficiently high to cause some ill defined banding for each increment as described earlier, but because the heat capacity is constant along the entire length of the calorimeter this effect will not be reflected in the shape of the heat versus coverage curve. Unfortunately, however, we cannot be certain that the initial rise in the heat curve is not the result of an artefact. The films were unavoidably patchy and because there was some lack of lateral uniformity in the heat capacity distribution, it is conceivable that the increase in heat has its origin in this trivial cause and that it is not the consequence of some intrinsic variation in the heat of adsorption with coverage.

As in the case of germanium, there appears to be a value of coverage, namely $\theta = 0{\cdot}7$, which has a special significance. It is in this region that the heat of adsorption begins to fall and it is also here that the film resistance ceases to increase with coverage. At $\theta = 0{\cdot}8$, the heat falls very rapidly and this coincides roughly with the point at which EISINGER and LAW[11] find the sticking coefficient also diminishes rapidly. We again note that our silicon surface is *n*-type, but the characteristics of the oxygen seem to be independent of whether the material is *n*- or *p*- type.[11] An attempt to interpret the slope of the heat curve meets with the same difficulties as for germanium, but it is probable that the explanation is essentially the same in each case.

Our finding of a coverage corresponding in effect to one atom of oxygen per surface silicon atom at the end of the fast adsorption confirms the usage of LAW[12]. It throws doubt on the validity of the use by SCHLIER and FARNSWORTH[10] of larger θ values at full coverage and therefore on the correctness of their calculated values of the sticking coefficients at intermediate coverages.

REFERENCES

1. GREEN M., KAFALAS J. A. and ROBINSON P. H., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 349. University of Pennsylvania Press, Philadelphia (1957).
2. BEECK O., *Advanc. Catalys.* **2,** 151 (1950).
3. BAGG J. and TOMPKINS F. C., *Trans. Faraday Soc.* **51,** 1071 (1955).
4. WHABA M. and KEMBALL C., *Trans. Faraday Soc.* **49,** 135 (1953).
5. RIDEAL E. K. and TRAPNELL B. M. W., *Proc. Roy. Soc.* A**205,** 409 (1951).
6. BREWER L., *Chem. Rev.* **52,** 1 (1953).
7. KLEMPERER D. F. and STONE F. S., *Proc. Roy. Soc.* A**243,** 375 (1957).
8. HAYWARD D. O., Ph.D. Thesis, Liverpool University (1958).
9. SCHLIER R. E. and FARNSWORTH H. E., *Advanc. Catalys.* **9,** 434 (1957).
10. SCHLIER R. E. and FARNSWORTH H. E., *J. Chem. Phys.* **30,** 917 (1959).
11. EISINGER J. and LAW J. T., *J. Chem. Phys.* **30,** 410 (1959).
12. LAW J. T., *J. Phys. Chem. Solids* **4,** 91 (1958).

VOL.
14
1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 124–130. Printed in Great Britain.

THE ION BOMBARDMENT, OXIDATION AND REGENERATION OF GERMANIUM SURFACES

S. P. WOLSKY and E. J. ZDANUK

Research Division, Raytheon Company, Waltham, Massachusetts

Abstract—A sensitive quartz microbalance and an omegatron type mass spectrometer were used in an ultra-high vacuum system to investigate the interaction of oxygen and germanium surfaces. Oxidation data at 3 mm pressure for ion-bombarded surfaces is discussed. The omegatron mass spectrometer provided valuable information concerning the nature and effect of experimental procedures upon the sample ambient. Such knowledge is necessary for (1) developing the proper experimental technique and (2) interpreting the experimental data. It is concluded that adsorption investigators must work with atomically clean surfaces of known crystal orientation and degree of perfection in order to obtain a more detailed knowledge of the oxygen–germanium interaction.

INTRODUCTION

THERE is general agreement that at pressures greater than 0·1 mm, the process of adsorption of oxygen on germanium surfaces consists of an instantaneous uptake followed by a slow further adsorption which proceeds according to a logarithmic rate law of the form $N = A + B \log t$ where N is the number of oxygen atoms adsorbed per cm^2 of surface in t minutes. The kinetic data obtained by different groups,[1–5] however, varies. The exact mechanism of the thermal restoration of oxygenated germanium surfaces is unknown. There is still considerable uncertainty, therefore, concerning the nature of the oxidation process.

This paper discusses the results of an investigation of oxygen adsorption on and regeneration of ion-bombarded germanium surfaces including one experiment in which the ambient composition was monitored almost continuously. The effect of experimental procedures upon the observed oxygen–germanium kinetics is also considered.

EXPERIMENTAL

A very sensitive quartz microbalance and an omegatron type mass spectrometer were combined in a system in which pressures of 10^{-10} mm were attainable. A detailed description of the apparatus and experimental method including the surface cleaning argon ion bombardment technique is available elsewhere.[1, 6–8] The combination of the microbalance and the omegatron provides information concerning changes both in the weight of the sample and the composition of the system ambient. In general weight changes of 0·15 μg could be detected. A discussion of the reproducibility of the microbalance measurements is presented below. The characteristics and application of the omegatron spectrometer have been discussed previously.[9, 10] The tube used in this work had a sensitivity of 5×10^{-10} mm with sufficient resolusion to completely separate masses forty and forty-four.

All experiments were with oriented single crystal samples of 0·002–0·004 in. thickness and total surface areas of 10–20 cm^2. The samples were heated with an external oven which fitted snugly around the quartz hangdown tube.

MICROBALANCE REPRODUCIBILITY

The results of an experiment in which the scatter of points was slightly greater than average, Fig. 1, will be used in a discussion of the reliability of the microbalance. The individual observed weight changes were limited by the minimal detectable quantity which was, in this instance, 0·14 μg or 3×10^{14} atoms/cm^2. The straight line which best fits the data, as determined by the least square method, has the equation $N = 8{\cdot}9 \times 10^{14} + 4{\cdot}1 \times 10^{14} \log_{10} t$. If a reasonable confidence level of 95 per cent is assumed, a range of slopes from 3·2 to

$5{\cdot}1 \times 10^{14}$ and of intercepts from 8·2 to 9·6 is obtained. The dotted lines in Fig. 1 are the possible plots at the 95 per cent level. Vibrational interference with the microbalance operation has rarely exerted a significant effect upon the experimental result. The procedure of making repeated readings of the balance reduces the error below the minimal detectable weight change.

RESULTS

The ambient compositions recorded at all feasible stages of an oxidation experiment are tabulated in Table 1. The analysis in column *A* is, as has been assumed, due chiefly to the desorption of argon.

(3) The argon adsorbed by the sample during the bombardment process is not readily completely removed. Although some of the argon detected during the 600° heating was undoubtedly desorbed from the quartz hangdown tube and the sputtered film, observations with an empty tube[11] which had been used in similar oxidation experiments have led us to conclude that most of the argon was released from the sample.

(4) Heating to 600° caused an increase in the CO and H_2O concentrations. These gases were probably

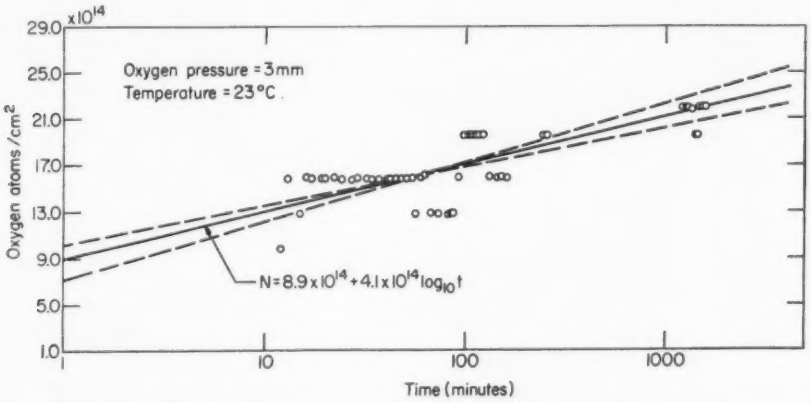


FIG. 1. The oxidation of germanium (microbalance reproducibility).

was that following a total of 52 hr of bakeout at 200–250°C. Since the omegatron is not usually used above pressures of 10^{-5} mm, the argon analysis was performed at a convenient pressure of 6×10^{-7} mm. The germanium sample was outgassed at 650–700°C for at least 8 hr prior to its use in this experiment. In considering Table 1, it is necessary to remember that both the sample and surrounding quartz tube were affected by such experimental procedures as ion-bombardment and heating.

The omegatron results reveal several interesting factors.

(1) The argon used in the bombardment was not a source of significant contamination of the system. Our gas handling technique and the argon analysis in column *B* have been discussed in detail elsewhere.[11]

(2) The large initial pressure rise generally observed on heating argon-ion-bombarded surfaces adsorbed by the quartz surfaces during such pre-experimental processing as cleaning and glass-blowing. Similar ambient changes occur during the bakeouts necessary to obtain ultra-high vacuum conditions, especially in Pyrex systems in which decomposition of the glass above 350°C releases water and other undesirable contaminants.

Following the second heating (column *H*), which is not a usual part of the experimental process, oxygen was admitted at 3 mm pressure. The kinetics, shown in Fig. 2(b), may be expressed by the equation $N = 3{\cdot}4 \times 10^{14} + 4{\cdot}9 \times 10^{14} \log_{10} t$. The sample was of (110) orientation. The interesting aspects of the kinetics are (i) the low initial uptake of ~0·4 monolayers of oxygen indicating partial contamination of the surface and (ii) the comparatively high rate of further adsorption so that after 1000 min the equivalent of 2·1 monolayers of oxygen had been formed. The sample

Table 1

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	
	Prior to experiment	After argon admitted to $6{\cdot}4\times10^{-7}$mm	After argon ion bombardment, 11 min at 700 V	Temp 25–180°C	Sample heating, temp 350–600°C	After 1·5 hr at 615°C	After pumping for 15 hr room temp	2nd heating cycle, after 2 hr at 615°C	Ionic species
Mass no.	Partial pressure (in mm Hg)								
>44	—	—	—	—	$1{\cdot}6\times10^{-9}$(74)[a]	$1{\cdot}2\times10^{-9}$(85)[a]	—	*T*(85–86)[a]	GeO^+, Ge^+
44	*T*	*T*	*T*	$2{\cdot}2\times10^{-9}$	$5{\cdot}6\times10^{-9}$	$3{\cdot}1\times10^{-9}$	—	$6{\cdot}0\times10^{-10}$	CO_2^+
40	—	$6{\cdot}2\times10^{-7}$	$2{\cdot}5\times10^{-8}$	$4{\cdot}6\times10^{-7}$	$1{\cdot}2\times10^{-6}$	$1{\cdot}0\times10^{-7}$	*T*	$1{\cdot}8\times10^{-8}$	A_{40}^+
28	$3{\cdot}6\times10^{-9}$	$8{\cdot}8\times10^{-9}$	$3{\cdot}5\times10^{-9}$	$2{\cdot}9\times10^{-8}$	$2{\cdot}2\times10^{-7}$	$1{\cdot}7\times10^{-7}$	$1{\cdot}9\times10^{-9}$	$1{\cdot}6\times10^{-8}$	CO^+, N_2^+
18	$6{\cdot}0\times10^{-10}$	$6{\cdot}0\times10^{-10}$	$6{\cdot}0\times10^{-10}$	$7{\cdot}8\times10^{-9}$	$4{\cdot}1\times10^{-8}$	$2{\cdot}2\times10^{-8}$	*T*	$7{\cdot}6\times10^{-9}$	A_{36}^+, H_2O^+
17[b]	*T*	*T*	*T*	*T*	$1{\cdot}1\times10^{-8}$	$5{\cdot}4\times10^{-9}$	—	$1{\cdot}6\times10^{-9}$	OH^+
16[d]	*T*	$7{\cdot}0\times10^{-10}$	*T*	$8{\cdot}4\times10^{-9}$	$1{\cdot}1\times10^{-8}$	$6{\cdot}1\times10^{-9}$	—	$8{\cdot}0\times10^{-10}$	CH_4^+, O^+, O_2^{++}
15[b]	*T*	*T*	*T*	*T*	$6{\cdot}7\times10^{-9}$	$3{\cdot}6\times10^{-9}$	—	*T*	CH_3^+
14[b]	*T*	*T*	*T*	*T*	$7{\cdot}8\times10^{-9}$	$5{\cdot}4\times10^{-9}$	—	$6{\cdot}0\times10^{-10}$	N^+, CH_2^+
12[b]	*T*	*T*	*T*	*T*	$6{\cdot}8\times10^{-9}$	$3{\cdot}6\times10^{-9}$	—	*T*	C^+
2[c]						$2{\cdot}0\times10^{-9}$		$6{\cdot}0\times10^{-10}$	H_2^+
Total pressure	$5{\cdot}0\times10^{-9}$	$6{\cdot}4\times10^{-7}$	$3{\cdot}0\times10^{-8}$	$6{\cdot}0\times10^{-7}$	$1{\cdot}4\times10^{-6}$	$3{\cdot}0\times10^{-7}$	$2{\cdot}1\times10^{-9}$	$4{\cdot}5\times10^{-8}$	

(a) Mass no.
(b) Mass of ions which represent fragments of molecules.
(c) Hydrogen monitored only in the instances shown.
(d) Mass 16 usually due predominantly to CH_4 in this experiment.

Operating Conditions: Accelerating voltage 50 V.
Electron collector voltage 22·5 V.
Trap voltage 0·52 V.
RF voltage 0·75 V.
Electron current 2 μA.

Ion gauge on continuously.

VOL
14
196

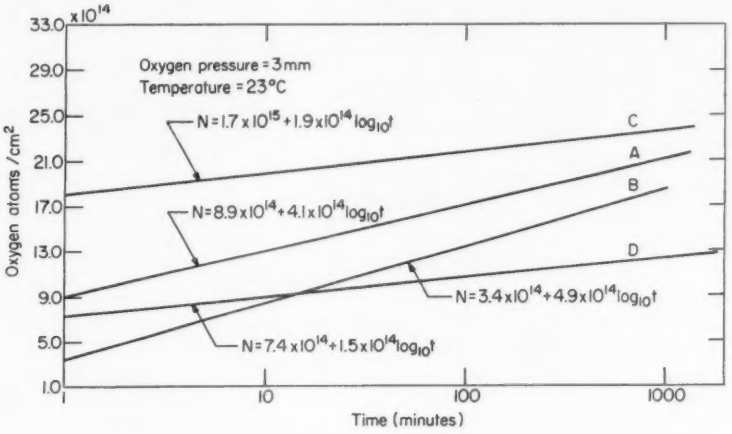


FIG. 2. The oxidation of germanium.

heating at 615°C for 2 hr prior to oxygen exposure should have regenerated any surface covered with GeO, it is likely, therefore, that the contaminant was GeO_2. Since the kinetics of the oxygen adsorption process as a function of orientation are not completely known, no quantitative comparison of the rate with other than (110) samples is valid. It has been suggested[12] that the rate of oxidation of the (110) plane is greater than that of the (111) or (100) in the range of room temperature to 550°C.

Plots *A*, *C* and *D* were obtained with (100) oriented samples. Samples *A* and *C* were given the usual experimental processing, whereas sample *D* remained in vacuum for 19 hr following bombardment and was annealed at only 400–425°C prior to oxygen exposure. The surfaces of *C* and *D* were similar in that both had been etched by the thermal and ion-bombardment treatment. Sample *A* showed no signs of etching. In general the samples are not etched by the experimental procedure. The equations for plots *A*, *C* and *D* are as follows:

$$\left.\begin{array}{ll}(A)\ N = 8{\cdot}9\times10^{14}+4{\cdot}1\times10^{14}\log_{10} t & \\ (C)\ N = 1{\cdot}7\times10^{15}+1{\cdot}9\times10^{14}\log_{10} t & (P = \\ (D)\ N = 7{\cdot}4\times10^{14}+1{\cdot}5\times10^{14}\log_{10} t & 3\text{ mm})\end{array}\right\}$$

A completely anomalous result is presented in Fig. 3. In this instance there appeared to be an induction period ($t<10$) when little or no adsorption occurred following which the process proceeded at a very high rate according to the equation $N = 0{\cdot}96\times10^{14}+5{\cdot}6\times10^{14}\log_{10} t$. The situation was further complicated by the observation of a film on the sample surface. A detailed investigation was not performed to determine the nature of this film. No weight loss was observed on removal of the oxygen by pumping.

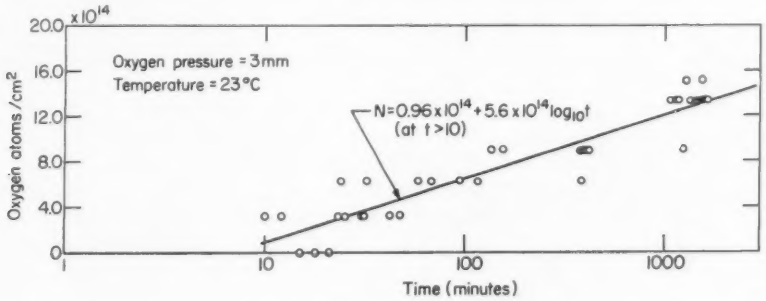


FIG. 3. The oxidation of germanium.

Table 2

(1)	$N = 8{\cdot}4\times10^{14}+1{\cdot}43\times10^{14}\log_{10} t$	$P = 0{\cdot}1$ mm[2]
(2)	$N = 7{\cdot}70\times10^{14}+1{\cdot}09\times10^{14}\log_{10} t$	$P = 66$ mm[3]
(3)	$N = 4{\cdot}60\times10^{14}+0{\cdot}50\times10^{14}\log_{10} t$ $N = 7{\cdot}42\times10^{14}+0{\cdot}70\times10^{14}\log_{10} t$ $N = 9{\cdot}65\times10^{14}+0{\cdot}89\times10^{14}\log_{10} t$	$P = 0{\cdot}1-0{\cdot}6$ mm[4]

DISCUSSION

The differences in the individual kinetic experiments described above are not unique. A brief survey (Table 2) of the adsorption data of others using comparatively large area powdered samples reveals similar disagreements.

These data were obtained with samples prepared by crushing in vacuum (equation 1), and chemical reduction in CO (equation 2) or H_2 (equation 3). In addition to the data in Table 2, other pertinent information has been reported. GREEN *et al.*[2] observed the uptake on crushed surfaces at a pressure of 60 mm to be $2{\cdot}4\pm0{\cdot}2$ layers of oxide after 18 min of exposure with no further detectable increase. LIGENZA[3] attributed similar results in his work to the catalyzing effect of moisture. SPARNAAY[5], using crushed surfaces in an oxygen pressure of 10 mm, noted a rapid uptake of 2–2½ oxygen atoms per surface germanium atom followed by a continuing slow reaction. Subsequent heating at 600°C caused regeneration of the original surface. Finally, the work on the heat of adsorption of oxygen on germanium films presented at this Conference[13] raises the question of the accessibility of oxygen to the surfaces in a tightly packed powder.

There appear to be several possible major mechanisms which either separately or in combination could produce the variable adsorption data observed.

(1) Partial contamination due either to insufficient initial surface cleaning or inadequate vacuum conditions would cause differences of the initial uptake in individual experiments. Whether the overall kinetic rate is affected may depend upon the nature of the contaminant.

(2) The presence of contaminants in the surface could, by having a catalytic effect upon the chemisorption process, affect the extent of immediate oxygen coverage as well as cause an increased overall rate.

(3) If the adsorption process is surface structure sensitive, as has been suggested, both the initial coverage and the subsequent rate would be affected.

It is apparent that several combinations of the above processes could explain the data of Fig. 2. The low initial uptakes are consistent with the existence of a partial contamination of the surfaces. The similarity in the rates of the etched surfaces *C* and *D* and their difference with that of sample *A* infers the possibility of surfaces which are energetically different because of structural dissimilarities. Of course, the catalyzing effect of a reactive contaminant could also explain the comparatively high adsorption rate observed with sample *A*. The anomalous oxidation, Fig. 3, in which an induction period occurred suggests a nucleation or rearrangement type process. No definite understanding of this experiment is available at this time. This result is offered mainly to illustrate further the large gap of uncertainty concerning the adsorption properties of germanium surfaces.

Although we have referred specifically only to our own data in the above discussion, it is felt that all reported oxidation data can be explained similarly. For example, surfaces prepared by vacuum bakeout or chemical reduction could readily show detailed differences both in surface contamination and structure. Further speculation on this subject is valueless until more detailed knowledge of the dependence of the oxygen adsorption on such factors is available.

The omegatron experiment has revealed information important to the surface investigations. A preliminary discussion of this data has appeared elsewhere.[10] The generation of both oxidizing

VOL. 14 1960

and reducing species by the experimental processing could cause competitive reactions at the sample surface. With the composition of the ambient dependent upon the history of the system, the cleanliness and physical state of the surface might be expected to vary in each experiment unless a very carefully planned rigorous and methodical treatment of the system is performed. Even after such processing, the movement of crushing or cleaving mechanisms may cause enough desorption from glass and metal parts to interfere in the measurements.

The comparatively large surface areas of the powdered samples should serve to mitigate serious surface contamination. Prolonged heating of the sample in uncertain ambients, or its repeated use in a series of experiments, however, could result in deleterious effects. The investigations of electrical properties, on the other hand, have generally involved samples of, at most, several square centimeters total surface area. The comparatively small area makes such measurements especially sensitive to even slight degrees of contamination. In most such experiments, fortunately, heating is accomplished by passing a current through the sample thereby avoiding any direct heating of the container walls. However, in clean surface work consideration should be given the possible effects of all experimental procedures. BILLS and EVETT[14], for example, have suggested that glass decomposition during bakeout, ion-bombardment or the heating of filaments near the container walls may contribute to contamination of a sample surface.

It is obvious that the vacuum technique is of prime importance. The use of a simple bakeable mass analyzing device such as an omegatron to monitor the ambient, therefore, is extremely valuable if not actually a necessity in the development of experimental technique.

The omegatron experiment indicated that very long post-bombardment heatings are necessary to completely remove the bombarding gas from the sample. This result is consistent with the observation of MADDEN and FARNSWORTH[15] who found the surface recombination velocity to be dependent upon the temperature and duration of heating. The effect of incomplete annealing procedures upon oxygen kinetics are not known but obviously if the adsorption is structure sensitive, careful control of the annealing procedure is necessary. A single experiment in this laboratory with a non-annealed surface showed a very high rate of oxidation. Further, more controlled experiments of the dependence of the adsorption on surface perfection are planned.

The ability to thermally regenerate oxygenated clean surfaces of germanium at temperatures $>500°C$ has been well established.[16–19] Partial regeneration has also been reported at 200°C[20] and other measurements[16, 21, 22] of surface properties have revealed effects which were reversed solely by prolonged pumping. It is very unlikely that one mechanism can be used to explain the restoration of clean surface properties over the temperature range of 25–600°C. The range from 400 to 600°C has been partially investigated in our laboratory[17, 18] and by workers using crushed samples.[19] The powder work,[19] however, was not definitive since (i) the nature of a surface subjected to repeated oxygenation and regeneration was assumed to be identical with that of the original crushed surface, and (ii) the analysis of the GeO film was subject to an error at least 20 per cent of the total oxygen involved. There is agreement amongst almost all investigators that the predominant mechanism of regeneration at 500–600°C is that of the evaporation of GeO from the surface. The thermal regeneration phenomenon has been investigated in a preliminary manner with the microbalance on the assumption that the observed weight loss on heating was due solely to the vaporization of GeO. This assumption appears valid since no significant pressure rises have been detected on heating as would be expected if oxygen, as such, was desorbed. On this basis it was found that ~12 per cent of the oxygen involved in the experiment could not be accounted for solely by the evaporation of GeO. The most reasonable mechanism, therefore, appears to be one involving the diffusion of oxygen into the bulk material. Such behavior has been observed for oxygen on both nickel and silicon surfaces. The diffusion into the sample in our experiment would have caused an increase in concentration of $\sim 1 \times 10^{17}$ oxygen atoms/air^3. Mass spectrometer experiments in this laboratory[23] have shown oxygen solubilities in germanium of the order of 10^{18} atoms/cm^3. It would not be necessary, therefore, to postulate any unique property for germanium. On the basis of our preliminary investigation it is proposed that

the thermal restoration of clean germanium surfaces at 500–600°C proceeds through the dual mechanism of the removal of oxygen from the surface as GeO and the diffusion of oxygen into the bulk germanium.

The room temperature reversible effects are not easily understood, especially since those studying adsorption have not reported desorption effects.[17, 19, 20] One possibility again, is the dependence of the quantity and mode of adsorption on the surface structure. More work in this area is obviously necessary.

CONCLUSION

From our discussion it appears that the main problem of adsorption investigators is still the preparation of atomically clean surfaces of known crystal orientation and degree of perfection. A detailed understanding of the nature of the interaction of oxygen and germanium will not be obtained until this problem is solved.

Acknowledgements—The authors wish to thank Mrs. P. M. RODRIGUEZ for her assistance in these experiments and Mr. F. HAWKES for his aid in the statistical analysis of the data.

REFERENCES

1. WOLSKY S. P., *J. Phys. Chem. Solids* **8,** 771 (1957).
2. GREEN M., KAFALAS J. and ROBINSON P. H., *Semiconductor Surface Physics*, p. 349. University of Pennsylvania Press, Philadelphia (1957).
3. LIGENZA J. R., *Bull. Amer. Phys. Soc.* ser. 2, **2,** 371 (1957).
4. DELL R. M., *J. Phys. Chem.* **61,** 1584 (1957).
5. SPARNAAY M. J., reported at *Brussels Conference on Semiconductors*, June 1958.
6. WOLSKY S. P. and FOWLER A. B., *Semiconductor Surface Physics*, p. 401. University of Pennsylvania Press, Philadelphia (1957).
7. WOLSKY S. P., *Phys. Rev.* **108,** 1131 (1957).
8. WOLSKY S. P., *Semiconductor Products* **2,** 36 (1959).
9. ZDANUK E. J., BIERIG R., RUBIN L. and WOLSKY S. P., 19*th Rep. Ann. Conf. Phys. Electronics*, p. 162. Massachusetts Institute of Technology, Cambridge, Mass. (1959).
10. WOLSKY S. P. and ZDANUK E. J., *Proceedings of the* 1959 *National Symposium on Vacuum Technology of the American Vacuum Society, Philadelphia* (1959).
11. WOLSKY S. P. and ZDANUK E. J., *Proceedings of the International Conference on Residual Gases in Electron Tubes and Associated Vacuum Systems, Lake Como* (1959).
12. LAW J. T. and MEIGS P. S., *Semiconductor Surface Physics*, p. 383. University of Pennsylvania Press (1957).
13. BRENNAN D., HAYWARD D. O. and TRAPNELL B. M. W., this Conference, preceding paper.
14. BILLS D. G. and EVETT A. A., *J. Appl. Phys.* **30,** 564 (1959).
15. MADDEN H. H. and FARNSWORTH H. E., *Phys. Rev.* **112,** 793 (1958).
16. SCHLIER R. E. and FARNSWORTH H. E., *Semiconductor Surface Physics*, p. 3. University of Pennsylvania Press, Philadelphia (1957).
17. WOLSKY S. P., *J. Appl. Phys.* **29,** 1132 (1958).
18. WOLSKY S. P. and ZDANUK E. J., *Bull. Amer. Phys. Soc.* ser. 2, **4,** 227 (1959).
19. ROSENBERG A. J., ROBINSON P. H. and GATOS H. C., *J. Appl. Phys.* **29,** 771 (1958).
20. LAW J. T. and GARRETT C. G. B., *J. Appl. Phys.* **27,** 656 (1956).
21. DILLON J. A. Jr. and FARNSWORTH H. E., *J. Appl. Phys.* **28,** 174 (1957).
22. HANDLER P., *Semiconductor Surface Physics*, p. 23. University of Pennsylvania Press, Philadelphia (1957).
23. PAPAZIAN H. A. and WOLSKY S. P., *J. Appl. Phys.* **27,** 1561 (1956).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 131–136. Printed in Great Britain.

THE MECHANISMS FOR SILICON OXIDATION IN STEAM AND OXYGEN*

J. R. LIGENZA and W. G. SPITZER

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

VOL. 14 1960

Abstract—This investigation is a study of some of the mechanisms which are important in the growth of SiO_2 films on high purity silicon surfaces. Films have been prepared by using O^{16} and O^{18} sources and their optical properties investigated by the use of infrared absorption measurements. Two of the fundamental infrared absorption bands show considerable isotopic shifts when films of $SiO_2{}^{16}$ and $SiO_2{}^{18}$ are compared. The SiO_2 films studied here, were prepared either by oxidizing silicon in oxygen at high temperature or by oxidation in high pressure steam at a somewhat lower temperature. Measurements of films prepared by successive isotopic oxidation show that in both the steam and oxygen cases considerable oxygen isotopic exchange has taken place which indicates that mobile oxygen species diffuse into the film from the silica–gas interface. In the steam oxidation, it is concluded that the mobile species is responsible for the growth of the film and the rate limiting step for the growth is at the silica–silicon interface. The results of the oxygen oxidation are not completely understood and while oxygen exchange is also observed in this case, further work is necessary to determine whether a diffusing oxygen species is responsible for the growth of the oxide film.

1. INTRODUCTION

THERE has been considerable effort[1] to understand the basic processes taking place during the growth of oxide films. The silicon–oxygen system is of great importance in device technology and in particular, heavy oxide layers may be used for passivating silicon devices.[2] The present study is an attempt to understand the mechanisms which play an important part in the growth of silica films. A number of amorphous SiO_2 films have been investigated. These films were produced through oxidation of high purity silicon by reacting with oxygen or high pressure steam at elevated temperatures. Oxygen 16 and 18 isotopes were used in the preparation of the films. Two of the infrared absorption bands of vitreous silica which were made with O^{18} show pronounced isotopic shifts when compared with films made with O^{16}. Several of the films used in this study were produced by oxidation in an atmosphere containing one type of oxygen isotope. The oxidation was then continued in an atmosphere containing the second isotope. The average isotopic oxygen concentration in the silica films was determined from measurements of the shift of the infrared absorption bands. The isotopic oxygen distribution was then obtained by measurements on films etched to different thicknesses.

The isotopic oxygen distribution furnishes some insight into the oxidation mechanism. For example, if the mechanism of oxide growth were by means of oxygen motion through the silica film, then all of the silica formed by using the second isotope would be at the silicon–oxide interface. Similarly, if a silicon ion, silicon ion vacancy or oxygen ion vacancy motion were predominant, then the concentration of the second isotope would be greatest at the oxide–gas interface. When isotopic oxygen exchange takes place then the distribution of the second isotope in the original film as well as in the newly grown film may be determined. Evidence for oxygen exchange in SiO_2 has been previously reported in the literature.[3]

2. METHOD OF ANALYSIS

It is well known that silica possesses several strong absorption bands in the spectral region between 5 μ and 30 μ.[4] These bands are attributed to the fundamental vibrations of silica. Curve (1)

* This work was supported in part by the Department of the Army under Contract DA36-039 sc-64618.

in Fig. 1 shows the transmission as a function of wavelength for a 3300 Å silica specimen. This specimen was produced by the oxidation of a polished silicon wafer. The oxidation was carried out by using pure oxygen 16 at atmospheric pressure and 950°C. The silicon was stripped from the film by a chlorine gas etch procedure described in Section 3(b). A preliminary analysis using classical dispersion theory[5] indicated that the optical constants, refractive index and extinction coefficient, which are obtained from this film, are compatible with those obtained from reflectivity measurements of bulk vitreous SiO_2 samples. Curve (3) in Fig. 1 shows the data for a sample produced by the oxidation of silicon in oxygen 16 steam at 120 atm pressure and 650°C. A comparison of curves (1) and (3) shows that the optical properties of the silica produced by the two different methods are the same. Curve (2) in Fig. 1 gives the data for the silica specimen which was made by using the steam oxidation method. However, in this case, the water which was used was 90 per cent D_2O^{18}. By comparing curve (2) to curves (1) and (3) it may be seen that two of the transmission minima in the O^{18} specimen have undergone appreciable isotopic shifts. The minima at 9·25 μ and 21·9 μ in the O^{18} specimens have shifted to 9·6 μ and 22·6 μ, respectively, in the 90 per cent O^{18} sample. A sample was also prepared by using water containing 50 per cent O^{16} and 50 per cent O^{18}. The data for the two absorption bands are given in Fig. 2. The observed minima lie midway between those for a pure O^{16} silica film and the expected position for a pure O^{18} film. Since a thin ($X<1$ mm) silicon wafer is quite transparent in both these spectral regions,[6] the results given here were obtained by measuring directly through the silicon as well as the silica

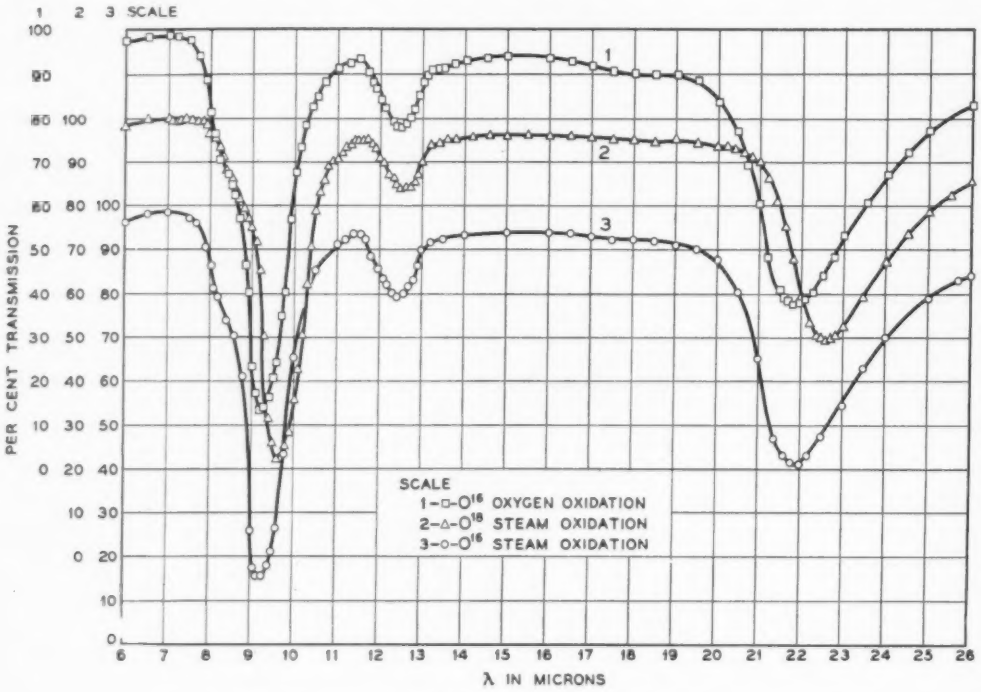


FIG. 1. Infrared transmission of three different SiO_2 films where the silicon has been removed. Film (1) was made by oxidizing silicon in oxygen 16. Films (2) and (3) were made by steam oxidations; (2) with D_2O^{18} and (3) with H_2O^{16}.

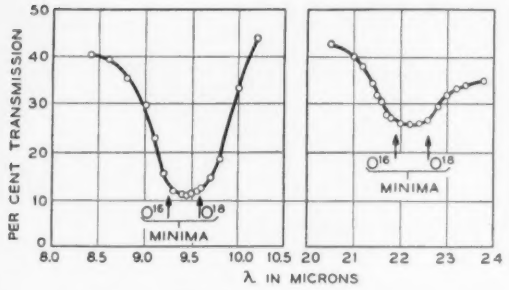


FIG. 2. Infrared transmission of a steam oxidized film where 50 per cent O^{16} and 50 per cent O^{18} water was used.

VOL. 14 1960

surface film. Fig. 3 shows plots of the oxygen 16 concentration divided by the total oxygen concentration versus wavelength of the transmission minima for each of the two absorption bands. As may be seen from this figure, the shift appears to be linear. It should be stressed that only the oxygen atoms which are chemically bonded to the silicon atoms are of importance in these measurements. Since we are measuring the fundamental resonance vibrations of the silica crystal, the oxygen concentration which does not participate in the formation of the SiO_2 crystal will not have

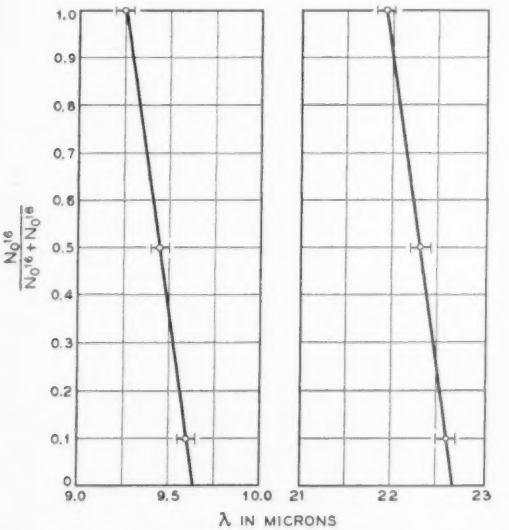


FIG. 3. A plot of the O^{16} concentration divided by the total $O^{16}+O^{18}$ concentration as function of wavelength for the two adsorption bands.

any pronounced effect on the strength or wavelength of the fundamentals. The results of the above measurements indicate that it is quite possible to measure the relative isotopic oxygen abundances in a silica film with an accuracy of ± 10 per cent.

3. METHODS AND MATERIALS

The infrared measurements were made with a single beam, double pass prism spectrometer. A sodium chloride prism was used to cover the wavelength region between 6 and 15 μ and a cesium bromide prism was used for wavelengths greater than 15 μ. The silicon samples used in these experiments were cut from floating zone, oxygen-free, high purity (resistivity $\sim 100\ \Omega$ cm) ingots. The specimens were 1 cm $\times$ 0·4 cm $\times$ 0·05 cm and the large area surfaces were optically polished. After oxidation, the sample was placed in the exit beam of the spectrometer and the transmisson measured by using the conventional sample in–sample out technique. The spectral half width of the beam at 9 μ is 0·05 μ and at 22 μ is 0·1 μ.

In those experiments where the wavelength shift of a transmission minimum was measured for several oxide thicknesses, a comparison sample which was not treated with the second isotope was taken through the measurement as well as the treated sample. Then the shift could be obtained from a direct comparison of the two samples. This procedure was adopted since the wavelength of the transmission minima may depend on the thickness of the oxide.[7]

(a) *Oxidation methods*

The oxidation in steam[8] was done in small metal bombs at 650°C and 120 atm of steam pressures. Before any oxidation, the silicon was boiled in concentrated nitric acid, boiled in distilled water and then dried. Under these conditions the oxidation rate, to form the amorphous oxide, is independent of time, directly proportional to the steam pressure and has an activation energy of 23 kcal/mole. The oxidation rates are strongly dependent on crystal orientation and for this reason oxidation rates are difficult to quote. The range of rates observed from randomly oriented silicon wafers is from 45 to 100 Å/min. Most specimens average 75 Å/min.

Ordinary water used in these oxidations was

purified by passing through a deionizing column followed by a double distillation in an all quartz system. D_2O^{18} water was obtained from the Isomet Corporation and was 90 per cent O^{18} and 1·7 per cent O^{17}. This water was vacuum distilled 3 times before use.

The oxygen oxidation was done in an open tube furnace of the globar type. The oxidations were carried out under 1 atm of oxygen and at a temperature of 950°C. The oxidation rate (2) to produce the amorphous SiO_2, under these conditions, is inversely proportional to the oxide thickness and directly proportional to the 4/5 power of the oxygen pressure. These rates can be approximately described by the equation

$$X^2 = 8{\cdot}4 \times 10^{10} p^{4/5} \left[\exp\left(\frac{-40{,}000}{RT}\right)\right],$$

where X is the film thickness in Å, t is time in minutes, p is oxygen pressure in atmospheres and R and T have their usual meanings. This equation holds over a pressure range of a few millimeters to 760 mm of Hg pressure and temperatures of 700–1100°C. The oxidation rate is only very slightly dependent on crystal orientation.

Oxide film thicknesses were determined by weight measurements with a microbalance and by measurement of transmission interference fringes with the infrared spectrometer. These methods agreed to within 5 per cent of each other.

(b) *Silicon stripping and film etching*

In some cases, the silicon substrate was stripped from the oxide film by etching[9] in chlorine at 800°C and at reduced pressures. The oxide is unaffected by the chlorine while the silicon is removed as volatile silicon chlorides. The etching geometry selected was such that the remaining silica window was supported by a frame of silicon.

To find the average concentration of converted oxide through the film thickness it was necessary to etch away portions of the oxide film. This was done by using a 2·6 molar HF solution. The etching rate was 350 Å/min at 27°C.

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

(a) *Steam oxidation*

Curve (1) in Fig. 4 shows the infrared data for a silica film which was produced in the following manner. The silicon was oxidized in oxygen 16 steam to a thickness of 2000 Å. The oxidation was then continued in oxygen 18 steam until the total oxide thickness was 3500 Å. Comparing curve (1) of Fig. 4 to the curves in Fig. 1 or to the plot in Fig. 3, we see that almost the entire 3500 Å film is composed of $SiO_2{}^{18}$. The film was then etched to 2000 Å, but the infrared absorption bands still showed the film to be almost completely $SiO_2{}^{18}$. A similar specimen was made except that the order of the oxidations was reversed. A 3000 Å oxygen 18 oxidation was followed by a 2500 Å oxygen 16 oxidation. In this case, the entire 5500 Å film is composed of $SiO_2{}^{16}$.

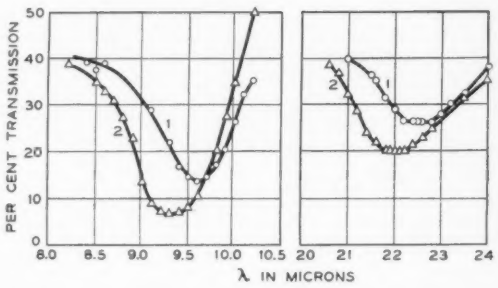


FIG. 4. Curve (1) is for a sample oxidized in O^{16} steam to 2000 Å followed by a 1500 Å steam O^{18} oxidation. Curve (2) is for a sample with a 2000 Å steam O^{18} oxidation followed by a 2000 Å O^{16} oxygen oxidation.

It is apparent that the exposure of an SiO_2 film containing one oxygen isotope to steam containing the second isotope results not only in the growth of new oxide but also in the conversion of the already existing film via an oxygen exchange mechanism. The diffusing species responsible for the exchange cannot be just silicon ions (or atoms), silicon vacancies, or oxygen vacancies. In order to accomplish the exchange, the oxygen species must diffuse into the film from the silica–steam interface. Since almost equal thicknesses of film were grown and converted via exchange, the exchange rate must be at least comparable to the growth rate. At the present time no experiments have been done to identify the diffusing species. It must be pointed out that these results do not rule out the presence of other mobile defects in the film.

In another experiment (Si-no.1), a 7700 Å $SiO_2{}^{16}$ film was produced on each of two silicon specimens. One of the two specimens was then treated

Table 1. Outline of the procedures used in obtaining samples 1A through 1F

Sample no.	Method of producing	Thickness (Å)	Treatment	Thickness after treatment (Å)
1*A*	oxygen 16 steam	7700		
1*B*	oxygen 16 steam	7700	oxygen 18 steam	7900
1*C*			etch 1*A*	5500
1*D*			etch 1*B*	5500
1*E*			etch 1*C*	3600
1*F*			etch 1*D*	3600

for 5 min in O^{18} steam at 650°C and 120 atm of pressure. Table 1 gives the sequence of treatments for Si-no.1; the infrared data are shown in a series of curves in Fig. 5. A comparison of the curves for samples 1*A* and 1*B* show a 30 per cent isotopic shift in the position of the minima. This shift, the result of a 5-min oxygen 18 steam treatment, increased the film thickness from 7700 Å to only 7900 Å. Therefore, the observed shift must be due, almost completely, to the exchange process. A comparison of curves (1*C*) and (1*D*) and curves (1*E*) and 1*F*) shows that the 30 per cent shift is a constant independent of film thickness. The fact that the exchange is constant with depth into the film implies that the concentration of diffusing species is uniform in the film for most of the 5 min used for the second stage of the oxidation. It is also implied that $\sqrt{Dt} \geq X$. If the time is assumed to be ~ 5 min, the thickness $X = 7700$ Å, then the diffusion constant D will have a lower limit of $\sim 2 \times 10^{-11}$ cm^2/sec. A different sample was oxidized and treated in the same manner as given for Si-no.1 except that the order of the isotopic oxidations was reversed. This sample yielded results which are identical with those given for Si-no.1. If the initial oxide film is grown by oxygen 16 oxidation and then treated by oxygen 18 steam for 5 min, the same result is obtained once again.

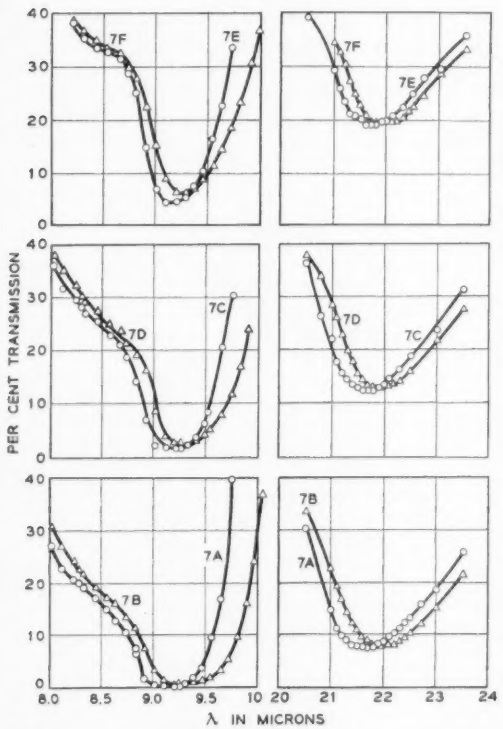


FIG. 5. Transmission as a function of wavelength for samples 1*A* through 1*F*. The samples are described in Table 1.

It appears probable that the actual oxidations as well as the rate limiting step for the oxidation takes place at the silicon–oxide interface. The 30 per cent exchange demonstrates that a high concentration of oxygen penetrates the film to the silicon–oxide interface and it, therefore, appears unlikely that silicon diffusion is at all necessary to explain the oxide growth. This is contrary to the results found in many other metal–oxide systems where the metal ion (or its vacancy) migration is the controlling mechanism for film growth.[1]

In order to demonstrate that the silicon substrate has little or no effect on the exchange process, a 12,000 Å silica film was made by oxidizing silicon in oxygen 18 steam. The silicon was then removed by etching (see Section 3b) and the infrared transmission measurements were taken (see curve 1 in

Fig. 6). The sample was then treated for 5 min in oxygen 16 steam. A comparison of curve (2) in Fig. 6, which represents the measurements taken after the oxygen 16 treatment, with curve (1) in the same figure shows a 30–35 per cent wavelength shift. This shift agrees with that observed for the sample of Fig. 5 where the silicon substrate was not removed.

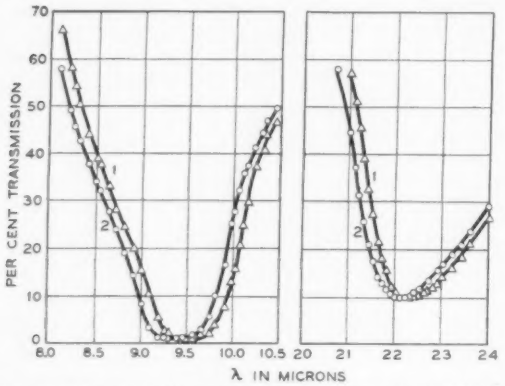


FIG. 6. Curve (1) shows the transmission of a 12,000 Å thick sample of SiO_2^{18} where the silicon has been removed. Curve (2) shows the same sample after a 5 min exposure to O^{16} steam at 650°C and 120 atm pressure.

(b) *Oxygen oxidation*

Curve (2) in Fig. 4 shows an example of almost complete conversion of an oxide layer where an oxygen oxidation was used. In this case, the initial oxygen 18 oxide layer of ~2000 Å was made by using the steam process. The oxidation was then continued to ≃4000 Å in an atmosphere of oxygen 16 at 950°C for 30 hr. As in the steam oxidation case, almost the entire film is composed of the isotope used in the second stage of oxidation. Therefore, the conclusions drawn from the first two steam oxidation experiments will also be true in the present case.

A number of experiments were done to determine the distribution of isotopic oxygen with film thickness, but results of these experiments are not well understood. The only certain conclusions to be drawn from these experiments are that the oxygen exchange rate at 950°C and 1 atm pressure is much slower than that with steam at 650°C and 120 atm, and that some type of oxygen species does diffuse through the oxide film during the oxidation. It is also true that in the oxygen oxidation, the rate for oxygen exchange and the rate for film growth are of similar magnitude. We do not observe any appreciable shift of the absorption bands except where there is significant oxidation. This behavior differs from that observed in the steam oxidation where large amounts of exchange were observed under conditions of very little film growth. In view of this difference and the difference in the growth rate dependence on film thickness, it is probable that the diffusing species for the two types of oxidation are not the same. The difference in the diffusing species is also indicated by the report[10] that after treating water free silica in water vapor at elevated temperatures (600–1200°C), an infrared absorption band is observed near 2·7 μ which has been attributed to the O–H stretching motion. It should also be pointed out that present experiments in the oxygen oxidation have not demonstrated whether or not the diffusing oxygen species are largely responsible for the growth of the oxide film.

Acknowledgements—We wish to express our appreciation to R. E. Cox for carrying out many of the oxidations, L. E. Howarth for his helpfulness in making a number of the infrared measurements and E. Tannenbaum for doing the chlorine etching. It is also a pleasure to thank M. M. Atalla and D. A. Kleinman for a number of stimulating discussions.

REFERENCES

1. Hauffe K., *Oxydation von Metallen und Metallegierungen*. Springer, Berlin (1956).
   Grimley T. B., *Chemistry of the Solid State* (edited by Garner W. E.) Academic Press, New York (1955).
   Kubaschewski O. K. and Hopkins B. E., *Oxidation of Metals and Alloys*. Academic Press, New York (1953).
2. Atalla M. M., *Proc. Conf. Elem. Comp. Semiconductors* (edited by Gatos H. *et al.*). Interscience Publishers, New York, to be published.
3. Hutchison D. A., *J. Chem. Phys.* **22,** 758 (1954).
4. Simon I. and McMahon H. O., *J. Chem. Phys.* **21,** 23 (1953).
   Reitzel J., *J. Chem. Phys.* **23,** 2407 (1955).
5. Spitzer W. G. and Kleinman D., to be published.
6. Collins R. J. and Fan H. Y., *Phys. Rev.* **93,** 674 (1954).
7. Born M. and Huang K., *Dynamical Theory of Crystal Lattices*, p. 123. Oxford University Press (1954).
8. Ligenza J. R., to be published.
9. Tannenbaum E., *J. Appl. Phys.*, to be published.
10. Moulson A. J. and Roberts J. P., *Nature, Lond.* **182,** 200 (1958).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 137–141. Printed in Great Britain.

SECTION IV. COMPOUND SEMICONDUCTOR SURFACES

RECENT WORK ON SURFACE PROPERTIES OF II–VI SEMICONDUCTORS

J. J. LANDER
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

VOL. 14 1960

RECENT work concerned with the semiconductor properties of single crystals of II–VI compounds has revealed several important surface effects. Zinc oxide has received by far the most attention, in part because its bulk properties are better known[1] and in part because its surface effects are often found to interfere with measurements of bulk properties. Therefore this work will be reviewed first, and work on the related materials CdS and CdSe will be mentioned later. Finally some of the problems encountered in using such data to interpret phenomena in sintered and thin film material will be discussed.

The surface charge and potential of a zinc oxide single crystal can be changed by chemical reaction with the environment, exposure to light in the near ultraviolet, application of a transverse electric field or bombarding with ions or electrons. The change in potential has been with respect to a bulk material which is *n*-type or near intrinsic, since material with significant *p*-type character has not been produced, nor have space charge layers with significant *p*-type character been produced by the adsorption of negative ions. The effects are properly attributed to surface conditions because they can be observed at room temperature and below, where diffusion of atomic species into the bulk is negligible. Equilibrium, however, is generally not attained rapidly at temperatures below about 300°C.

The surface chemical reaction between zinc oxide and its own atmosphere of zinc and oxygen is of fundamental importance in reaction systems involving zinc oxide. Analogous to treatments of imperfections in the bulk one might define intrinsic surface imperfections such as zinc vacancies and excess zinc atoms, noting, however, that this would be an oversimplification unless one took into account the various types of occupied and unoccupied sites. Instead, without further definition, we will denote the density of surface acceptor atoms as [O(s)] indicating that they are oxygen atoms, and the density of surface donor atoms as [Zn(s)]. The equations for the evaporation of oxygen and zinc, are

$$2\,O(s) \rightleftarrows O_2(g) \qquad (1)$$

$$Zn(s) \rightleftarrows Zn(g) \qquad (2)$$

and since the product of zinc and the square root of the oxygen vapor pressure is a function of temperature only, one also has

$$[O(s)]^2[Zn(s)] = K_3 \qquad (3)$$

in which K_3 is known. In the case of ZnO the oxygen to zinc pressure ratio can be changed through an extremely wide range. For example, at 300°C a pressure of oxygen of 1 atm implies an equilibrium pressure of zinc of about 10^{-40} atm.

The surface atoms may be considered as mostly neutral, but with ionization occurring at some according to the reactions

$$O(s) + e^- \rightleftarrows O^-(s) \quad K_4 = [O^-(s)]/[O(s)][e^-] \qquad (4)$$

$$Zn(s) \rightleftarrows Zn^+(s) + e^-$$

$$K_5 = [Zn^{+1}(s)]\,[e^-]/[Zn(s)] \qquad (5)$$

In non-equilibrium cases any one of the reactions (1), (2), (4) and (5) may be rate limiting, and more complicated situations are to be expected.

Reversible data for the surface conductivity of this system has been reported by THOMAS and LANDER[2] for zinc pressures in the range from 1 to 10^{-8} mm of Hg and temperatures between 300° and 500°C and they are reproduced in Fig. 1. The saturation value of $2\times10^{-4}\,\Omega^{-1}$/square corresponds to about 2 per cent of a monolayer of and values for the energies of ionization species are not known.

HEILAND[3, 4] has reported that hydrogen produces large and similar surface effects, but in that work the surface concentrations of hydrogen and excess zinc resulting from reduction of zinc oxide are not known. More generally it is believed that

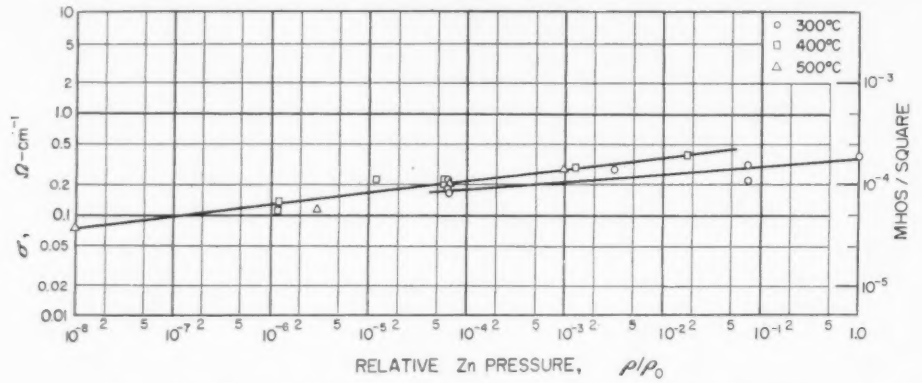


FIG. 1. Conductivity of 0·001 in. crystal at various temperatures as a function of zinc relative pressure.

positive charge compensated by electrons with average mobility taken equal to 100 cm²/V sec in a space charge layer beneath. Because the quantity adsorbed is not strongly dependent on zinc pressure the heat of adsorption of excess zinc must be high. Neglecting the degree of ionization of the excess zinc, an estimated value is about 10 kcal/mole for the concentration range explored. One hopes that further work with such tools as the field emission microscope, or slow electron diffraction would resolve the ambiguities in the model presented above.

Data for the temperature dependence of conductivity are given in Fig. 2 for a succession of "frozen-in" states. Similar results were obtained by evaporating minute quantities of zinc on to a crystal held in vacuum at room temperature. With increasing conductivity (increasing concentration of excess zinc) the slopes decrease and at the highest conductivities there is little temperature dependence in the range to liquid helium. This is presumably the result of overlapping donor orbitals. Because of the probable presence of traps at and near the surface the significance of the limiting slope at low concentrations has not been explained, strong reducing agents produce excess positive charge and strong oxidizing agents excess negative

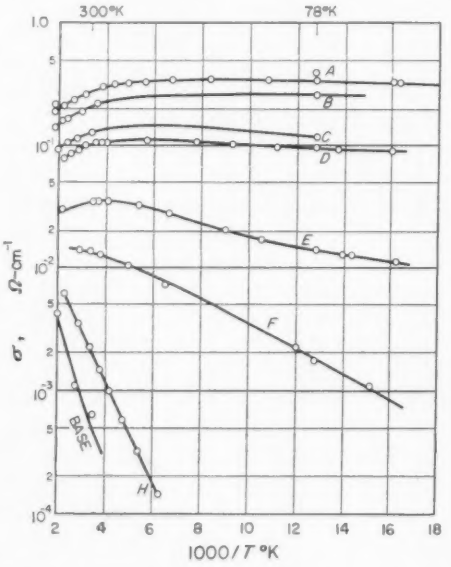


FIG. 2. 0·0019 in. diameter crystal conductivity as a function of temperature for different zinc coverages.

VOL. 14 1960

charge on zinc oxide surfaces, but chemical reduction or oxidation need not take place. Thus Thomas finds that cesium evaporated onto ZnO at room temperature behaves much like zinc.[5]

Large increases in conductivity of zinc oxide can also be obtained by exposure to ultraviolet light.[4] The results are dependent on ambient as well as on the intensity and frequency of the light. Since it was found that an increased conductivity produced in a crystal in vacuum at room temperature in this way is permanent, but could be reversed by exposure to oxygen, COLLINS and THOMAS[6] concluded that light results in a desorption of lattice oxygen atoms, a process which may properly be called photolysis. Changes in conductivity due to trapping processes unaccompanied by oxygen desorption are also possible and in particular it was observed that after a change in conductivity produced at liquid nitrogen temperature the initial state was regained at about $-150°C$ by warming the crystal in the dark. Presumably the oxygen does not desorb at very low temperatures. Data obtained by KRUSEMEYER[8] show that the efficiency of the process (the number of surface charges produced per absorbed photon) is near unity when the surface is far from saturation, and with sufficiently intense light saturation values of conductivity correspond to about 1 per cent of a monolayer of charge. Rates of change of conductivity, which are a function of the ambient, were also reported. Furthermore, surface photoconductivity can be produced by light with a frequency in the exciton range[7] (about 3·3 eV in zinc oxide). Earlier work on photoconductivity in thin films and sintered samples has been reviewed by HEILAND *et al.*[1]

Field effect data as a function of surface conductivity have been obtained at 70°K by HEILAND[4] and as a function of contact potential and surface conductivity at room temperature by KRUSEMEYER[8], who observed a range of contact potential from about $-0·1$ V to about $+0·5$ V with respect to the bulk potential (Fig. 3). In this range field effect mobility of different crystals varied from about 1 to about 100 cm^2/V sec, and it was concluded from the nature of the variation of conductivity with contact potential that the low values were caused by surface states rather than by screening of the field by ions in the space charge region. Application of the electric field to crystals in vacuum produced slow changes in surface potential. KRUSEMEYER also reported a quantitative treatment of the hole-trapping process which is in good agreement with the experimental results.

TAKAISHI[9], and KRUSEMEYER and THOMAS[10] have analyzed the equilibrium adsorption and charge transfer process at semiconductor surfaces (represented for ZnO by equations 1 to 5) as a function of partial pressure and composition of reacting vapor, and GARRETT has extended this type of analysis to a treatment of the absolute rates of catalytic reactions involving charge transfer in semiconductors.[11]

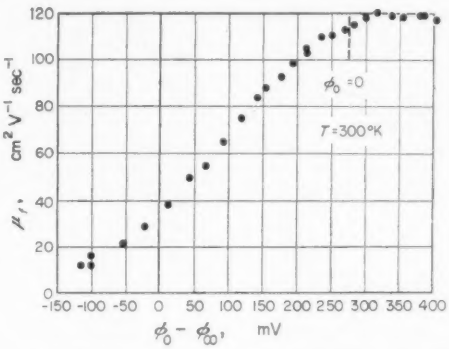


FIG. 3. Field effect mobility versus surface potential.

DEWALD[12] has measured the capacitance of single crystal zinc oxide electrodes in contact with aqueous electrolytes. He observed no surface-state effects over a wide range of bias (about 2 V) and bulk electron density (10^{14}–10^{19} electrons/cm^3). Variable surface dipoles and effects of low-lying donors (e.g. boron at 0·3 eV) were observed.

Work on surface effects in CdO and ZnS single crystals has not been reported but effects have been observed in the photoconductivity response of CdS and CdSe single crystals. BUBE[13] has reported a study of the "photostimulated, temperature dependent, reversible sorption of oxygen" resulting from treatment of (*n*-type) CdSe crystals at 300°C for 3 hr in vacuum which apparently left excess cadmium near the surface (a result similar to the *A* and *B* type surface effects in zinc oxide reported by HEILAND[1]). Subsequently conductivity was observed to decrease by as much as 5 orders of magnitude with exposure to oxygen at

lower temperatures and this was accompanied by an increased sensitivity of photoresponse (Fig. 4). Similar effects were observed by BUBE[14] in work with CdS single crystals. As a matter of practical technology one notes that there is a variable fraction of photoconductive response in these

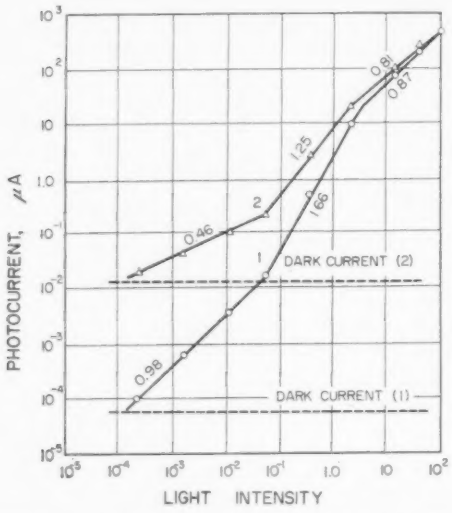


FIG. 4. Variation of photocurrent (for 1·5 V) as a function of light intensity at 75°C for crystal in (1) low-conductivity state, and (2) high-conductivity state. Light intensity of 100 corresponds to 900 ft-cd.

materials which depends on the state of the surface and which can be altered by temperature, atmosphere, incident light and electric field.

In summary, study of the surface properties of single crystals of II–VI semiconducting compounds exhibits large effects due to non-stoichiometry. Thus, in comparison with group IV elements, a more complex intrinsic chemistry must be taken into account. However, one does not have to worry so much about surface oxide films, at least on oxides. The relation between contact potential and conductivity appears to be better understood, and some surface trapping and recombination effects observed are very different (in zinc oxide recombination effects appear to be negligible). Many experimental techniques and much of the theory developed in work on germanium and silicon is readily adaptable to these materials, but, even in the case of zinc oxide, much remains to be done. Thus one might consider use of low voltage electron diffraction and the field emission microscope to obtain improved models of surface structure, analysis of combined measurements of surface conductivity, Hall coefficient and magnetoresistance as developed by ZEMEL and PETRITZ with germanium,[15] the surface p.e.m. effect, infrared absorption and conductivity, space charge effects on drift mobility as developed by HARRICK[16] (but not with zinc oxide where the hole life-time is presumably too short), electron spin resonance and extensive studies of chemical reactions at surfaces.

The remainder of this review will be concerned briefly with properties of II–VI compounds in sintered or thin film form. These have practical application in catalysis, photodevices, luminescence, thermionic emission and transistor-like devices. The importance of surface properties is much greater in such material and space charge effects at intergranular boundaries or across individual grains may dominate the semiconducting properties. HUTSON[17] has reviewed the problem of interpreting the results of experiments. In general, detailed knowledge of the bulk and surface chemical and physical properties as well as the granular geometry is required in order to arrive at quantitative predictions. However, some properties may be dominated by surface effects through a wide range of grain size. SORROWS[18] has reviewed work on thin films of lead sulfide and reported results of a study in which field effect and photoresponse were observed to have identical time constants. Thus the rate of transfer of charge into majority carrier traps (presumably at the surface) was the rate limiting process for both effects, and the density of charge of the traps determined the photoconductivity. The variation of photoresponse was correlated with trap density which could be changed by oxygen treatment. Further work will be reported by ZEMEL and VARELA[19].

A note of caution should be raised regarding the interpretation of data for material in sintered or thin film form. In some work it is likely that a simple model of the surface and intergranular regions may not be appropriate because of the precipitation during cooling of one or more insoluble phases. For example, several studies have been reported of the semiconducting and catalytic

properties of zinc oxide containing various concentrations in the range from 0·1 to 2 mol% of impurities such as lithium on one hand and indium or gallium on the other which respectively reduce and enhance conductivity.[1, 20] In one such study by RUDOLPH[21] *p*-type conductivity was reported for a material doped with 2 per cent lithium.[21] However, it is known that although lithium and indium oxide are moderately soluble in zinc oxide at high temperatures, insoluble compounds tend to precipitate on cooling. Thus THOMAS[22] found that the solubility of indium in zinc oxide ranges from 3×10^{19} atom/cm^3 at 1300°C to about 1×10^{17} at 800°C and he has studied the precipitation on dislocations at lower temperatures in single crystals. The precipitated phase was not identified but it is known that at least three phases exist in the ZnO–In_2O_3 system. Similar results have been obtained with lithium,[23] which, however, has a somewhat higher solubility, but diffuses at a lower temperature.

PIPER and WILLIAMS[24] review the properties of sintered electroluminescent materials (ZnS dominates the field) and the basic mechanisms of electroluminescence. The acceleration of conduction electrons to optical energies in regions of high field is the excitation process. These regions occur at grain boundaries and, similar to surface regions, their properties are no doubt sensitive to grain boundary composition, which may be readily affected by the ambient at moderate temperatures.

Finally, it should be noted that surface effects may dominate the behavior of sintered material and thin films even at high temperatures. DEREN *et al.*[25] have discussed the temperature dependence of conductivity of zinc oxide in the range to 700°C in terms of the equilibria for oxygen chemisorption. The range in which reversible effects are observed is a function of sintering temperature and ends at about 400°C for material sintered below 1000°C. They conclude that conductivity in material sintered at temperatures much higher than 1000°C is determined by bulk properties.

REFERENCES

1. HEILAND G., MOLLWO E. and STÖCKMANN F., *Solid State Physics* **8,** 191 (1959).
2. THOMAS D. G. and LANDER J. J., *J. Phys Chem. Solids* **2,** 318 (1957).
3. HEILAND G., *Z. Phys.* **138,** 459 (1954); **142,** 415 (1955); **148,** 15 (1957).
4. HEILAND G., *J. Phys. Chem. Solids* **6,** 155 (1958).
5. Private communication.
6. COLLINS R. J. and THOMAS D. G., *Phys. Rev.* **112,** 388 (1958).
7. POWERS M. and THOMAS D. G., to be published.
8. KRUSMEYER H. J., *Phys. Rev.* **114,** 655 (1959).
9. TAKAISHI T., *Leitfaden für Naturforscher* **11,** 286. (1956).
10. KRUSMEYER H. J. and THOMAS D. G., *J. Phys. Chem. Solids* **1–2,** 78 (1958).
11. GARRETT C. G. B., to be published.
12. DEWALD J. F., reporting at this Conference.
13. BUBE R. H., *J. Chem. Phys.* **27,** 496 (1957).
14. BUBE R. H., *Phys. Rev.* **101,** 1668 (1956).
15. ZEMEL J. N. and PETRITZ R. L., *Phys. Rev.* **110,** 1263 (1958).
16. HARRICK N. J., *Phys. Rev. Letters* **2,** 199 (1959).
17. *Semiconductor Chemistry* (edited by HANNAY N. B.) p. 542. *Monogr. Ser. Amer. Chem. Soc.*, Reinhold, New York (1959).
18. SORROWS H. E., Ph.D. Thesis, Cathodic University of America (1958).
19. ZEMEL J. and VARELA J., this Conference, *J. Phys. Chem. Solids* **14,** 142 (1960).
20. CVETANOVIC R. J. and FOSTER N. F., *Disc. Faraday Soc.*, to be published (1959).
21. RUDOLPH J., *Z. Naturf.* **14**a, 736 (1959).
22. THOMAS D. G., *J. Phys. Chem. Solids* **9,** 31 (1959).
23. LANDER J. J., to be published.
24. PIPER W. W. and WILLIAMS F. E., *Solid State Physics* **6,** 95 (1958).
25. DEREN J., HABER J. and WILKOWA T., *Z. Phys.* **155,** 453 (1959).

VOL. 14 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 142–149. Printed in Great Britain.

SURFACE PROPERTIES OF PbS PHOTOCONDUCTORS

J. N. ZEMEL and J. O. VARELA

U.S. Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland

Abstract—A study of the field effect trapping time constant as a function of temperature has been carried out on a series of lead sulphide photodetectors. The results show that the density of field effect traps (or surface traps) does not vary appreciably but that the carrier concentration increases as the sensitivity of the films decreases. The mobilities obtained from the high frequency field effect measurement agree quite well with Hall data on similar films. The rate of capture of the holes (the majority carriers) was relatively independent of the films whereas the rate of emission of holes from the surface levels increased with increasing sensitivity. A simple Shockley–Read model does not explain the various phenomena adequately. However, the general picture of majority carrier photoconductivity is substantiated by these measurements.

INTRODUCTION

THIN films of semiconductors have a unique advantage when surface properties are studied. The surface area to bulk volume ratio can be made sufficiently large so that bulk effects are minimal if not non-existent. While this aspect is advantageous, the problem of preparation is quite difficult. The lack of uniformity in evaporated films is well known to workers in the field and the sensitivity to ambients is critical for certain electrical properties. In recent years, thin films have been grown in a more or less reproducible fashion by epitaxial growth and chemical deposition. The former method is used when single crystals are desired; while the latter is used when economical production methods are important. An excellent example of chemically deposited films is the PbS photodetector.

PbS is one of the "older" semiconductors in that its non-metallic behavior has been known for years. Its greatest use recently has been as a detector of infrared radiation. The mechanism responsible for its sensitivity to radiation has been a subject of intensive study and almost equally intensive controversy.[1–3] However, our interest in this material is due to the fact that films of particular properties can be prepared at will and that the surface appears to dominate the film properties. The experiments suggest that many of the apparent irregularities in time constant, resistance, etc., are due to changes in trap densities.

A brief history of lead salt photoconductivity is pertinent in that it shows the strong dependence of film properties on surface treatment. Prior to the development of chemically deposited films, the lead salt films were evaporated on a glass substrate and then exposed to small doses of gaseous oxygen until maximum photosensitivity was obtained.[4] HUMPHREY and SCANLON[5] and others showed conclusively that only oxygen would induce photosensitivity in PbSe. In these experiments, electron accepting elements such as the halogens and sulfur changed the Fermi level quite drastically, driving the originally *n*-type films to *p*-type. However, only oxygen produced both a change in resistance and photoconductivity life time. This implied that only oxygen introduced a state which could produce photoconductivity. The implication of oxygen as the sensitizing element was sufficiently well established that PETRITZ *et al.*[6] could turn to the mechanism of photoconductivity. They argued that only surface traps could account for the estimated trap density assuming crystallite sizes reported in the literature. ZEMEL[7] showed the presence of "slow" surface states on these films. SORROWS[8] then showed conclusively the presence of "fast" surface states on PbS. The significant fact was the demonstration that the field effect trapping time constant and the photoconductive time constant were equal over a wide range of film conditions and d.c. light bias. This result implied that the photoconductive process

VOL. 14 1960

was limited by the trapping of holes as required by PETRITZ *et al.* theory.[6] The recombination of holes at this site is suggested by the work of PETRITZ and LUMMIS[9] on generation-recombination noise.

In this paper, essentially the same measurements done by SORROWS[8] are performed. The aim is to elucidate the trapping mechanism by means of the temperature dependence of the time constants. A Shockley–Read process[10] is invoked to account for the trapping, with moderate success.

EXPERIMENTAL METHODS

SORROWS[8] has given a detailed account of the experimental scheme. We shall only indicate it schematically. The circuit is shown in Fig. 1. An

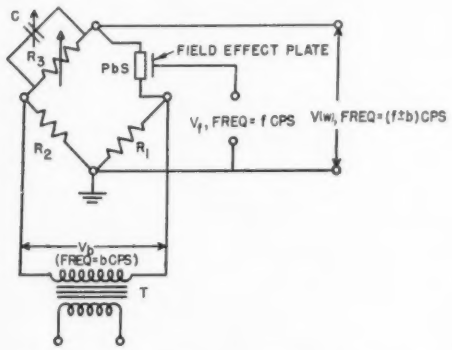


FIG. 1. Beat frequency bridge arrangement for measuring field and photo effects.

a.c. bridge, consisting of a variable resistor, PbS film and the balanced output windings of an audio oscillator, is balanced at a frequency b. A field effect voltage of frequency f is applied to the capacitor plate and the magnitude of the induced resistance change is measured as a voltage, $V(f \pm b)$ on the wave analyzer as a function of f. This has the desirable effect of virtually eliminating from the field effect measurement the dependence on the resistance of the films. Drifts will not cause difficulties and the a.c. current will minimize polarization of the film. A typical run is shown in Fig. 2 including a correction for the loading of the external circuit on the film.

A series of seven films was studied whose time constants and photoresponses at room temperature varied by a factor of 6. The films were maintained in a dry nitrogen ambient throughout the measurements and were stored under nitrogen in a vacuum dessicator. For measurements, the samples were placed in a sealed container under a dry nitrogen atmosphere. The physical arrangement is shown in Fig. 3. A d.c. light bias was provided by a 6-V lamp in a separate attached chamber. Photoconductive time constants could be measured using the experimental apparatus and modulating the resistance with a neon bulb driven by a Tektronix 105 square wave generator.

The light bias provided by the lamp operated from a well filtered d.c. constant voltage source. Changes in conductance of up to 100 per cent were made and the time constants were examined for various background illuminations.

The temperature dependent measurements were carried out over a range from $-45°C$ to $+45°C$. The temperatures were stabilized by immersing the container in constant temperature baths. At temperatures above 15°C a bath with a stability of better than $\pm 0{\cdot}005°C$ was used. Below 15°C, a special refrigerator was used which had a stability of $\pm 0{\cdot}25°C$. Reproducibility of the time constants was usually of the order of 10 per cent.

The field effect voltage induces a charge $Q_0 \exp(i\omega t)$ into the film where $\omega = 2\pi f$. The free charge changes the sheet conductivity, $\Delta\sigma$, of the film so that

$$\Delta\sigma = (\sigma - \sigma_0)d = |\Delta\sigma| \exp(i\omega t)$$
$$= |\Delta p| q\mu^* \exp(i\omega t) \qquad (1)$$

where σ is the instantaneous conductivity of the film, σ_0 the zero field conductivity, d is the film thickness, $|\Delta p|$ is the maximum induced charge and μ^* is the mobility of carriers in the film. One can easily show that the change in film resistance R is

$$\Delta R = \frac{W}{L} R^2 \Delta\sigma \qquad (2)$$

where W and L are the width and length of the capacitor plate, respectively. The change in voltage across the sample is due to the mixing of the bridge current $i_0 \exp(i\omega_b t)$ and the field induced resistance change. Therefore

$$V_{\Delta R} = \frac{W}{L} i_0 R^2 |\Delta\sigma| \exp[i(\omega \pm \omega_b)t] \qquad (3)$$

The bridge bias voltage, V_b, is $2i_0R$ since the cell resistance is the limiting resistance in the circuit.

Due to the capacitive and resistive loading of the cell, a correction factor is used as described by frequency generator in series with the film and measuring the output voltage with the wave analyzer at frequencies $f \pm b$. The ratio of the output $V_{\text{comp}}(f \pm b)$ to V_b is $L(\omega)$. Using this definition

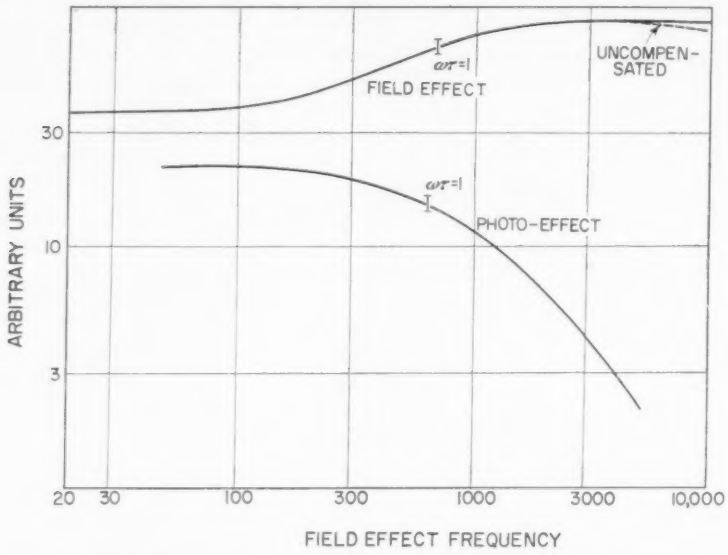


FIG. 2. Beat frequency voltage and photo response versus frequency for film 118–8 at 21°C in a dry nitrogen ambient.

SORROWS[8]. The measured r.m.s. signal on the wave analyzer is then

$$V_{\text{WA}} = \frac{W}{L} \frac{V_b L(\omega) R}{2\sqrt{2}} |\Delta p| \mu^* q \qquad (4)$$

where equation (1) is used for $|\Delta\sigma|$. The correction for the circuit loading of the film, $L(\omega)$, is obtained by inserting a constant voltage, variable

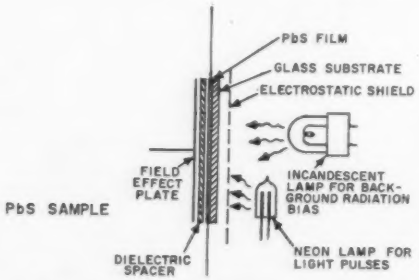


FIG. 3. Diagrammatic arrangement of PbS film, field effect electrode and light sources in container.

of $L(\omega)$ we find

$$V_{\text{WA}} = \frac{W}{L} \frac{V_{\text{comp}}(f \pm b) R}{2\sqrt{2}} |\Delta p| \mu^* q \qquad (5)$$

The excess free carriers can be calculated from a rate equation so that

$$q|\Delta p| = Q_0 \frac{[(1-\alpha+\omega^2\tau^2)^2 + \alpha^2\omega^2\tau^2]^{1/2}}{1+\omega^2\tau^2}$$

$$= Q_0 f(\alpha,\omega,\tau) \qquad (6)$$

where τ is the field effect time constant and α is the fraction of induced charge trapped in surface states at frequency $\omega \ll 1/\tau$.

Introducing equation (6) into equation (5) leads to

$$V_{\text{WA}} = \frac{W}{L} \frac{V_{\text{comp}}(f \pm b) R}{2\sqrt{2}} Q_0 \mu^* f(\omega,\tau,\alpha) \qquad (7)$$

This, in fact, is the form found for the experimental data.

VOL. 14 1960

At sufficiently high frequencies, i.e. $\omega \gg 1/2\pi\tau$, $(\alpha,\omega,\tau)\rightarrow 1$. Then the mobility is simply obtained from equation (7). Once the mobility is known, the carrier concentration and trap density can be obtained.

One makes use of two auxiliary definitions to define the rate of hole capture to and hole emission from the surface states. They are: $\tau_{12} = \tau/\alpha$, time of capture; $\tau_{21} = \tau/(1-\alpha)$, time of emission. Our interest will center on τ_{12} and τ_{21}.

EXPERIMENTAL RESULTS

In Table 1, the room temperature characteristics of the cells studied are listed. In Figs. 4 and 5 the dependence of $1/\tau_{12}$ and $1/\tau_{21}$ on reciprocal temperature is shown. The data strongly suggest an exponential behavior of the type

$$\frac{1}{\tau_{12}} \sim \exp(-\Delta E_{12}/KT)$$

$$\frac{1}{\tau_{21}} \sim \exp(-\Delta E_{21}/KT)$$

The values of ΔE_{12} and ΔE_{21} are listed in Table 2. The values of ΔE_{12} were quite close to ΔE_{21} for the more sensitive films. As the carrier concentration increased, ΔE_{21} remained substantially constant up to a carrier concentration of approximately $10^{16}/\text{cm}^3$. Above that, ΔE_{21} appears to decrease somewhat from its previous value of 0·23 eV. Similarly, there appears to be a decrease in ΔE_{12} for the least sensitive films. The ΔE_{12} for 118–18 appears to be anomalously low. The intercepts of $1/\tau_{21}$ at $1/T = 0$ appear to increase with increasing carrier concentration whereas $1/\tau_{12}$ is quite insensitive to variations in carrier density. The explanation of this behavior will be central to any theoretical explanation.

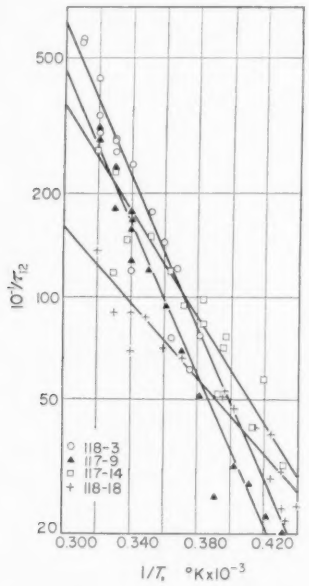


FIG. 4. $1/\tau_{12}$ vs. reciprocal absolute temperature for various representative films.

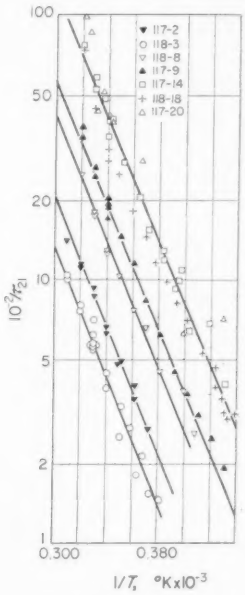


FIG. 5. $1/\tau_{21}$ vs. reciprocal absolute temperature.

As will be shown in the next section, the density of occupied surface states is

$$N_{ts} = \frac{\alpha}{1-\alpha} pd \qquad (8)$$

where pd is the product of the carrier concentration and the cell thickness. pd is proportional to $1/R$. In Fig. 6, $(1-\alpha)/\alpha$ is plotted against conductance for one film where R has been modulated with a d.c. light bias. As can be seen there is a

Table 1. Room temperature (21°C) characteristics of PbS films

Cell	Resistance ($\times 10^{-3}$ Ω)	Time constant ($\times 10^{+6}$ sec)	α	μ^* (cm²/V sec)	$p \times 10^{-15}$/cm³	$d \times 10^4$ cm
117–20	63	133	0·336	—	74	0·8±0·2
118–18	76	265	0·200	3·1	74	0·0±0·2
117–14	100	194	0·324	5·0	37	0·8±0·2
117–9	242	270	0·447	6·2	13	0·8±0·2
118–8	371	312	0·597	9·0	5·4	0·8±0·2
118–3	866	350	0·845	5·6	4·0	0·8±0·2
117–2	884	362	0·775	5·3	4·0	0·8±0·2

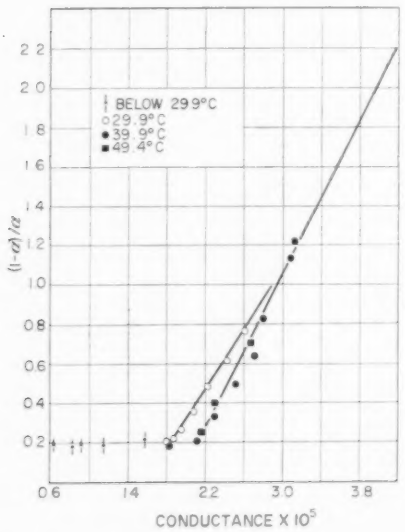


FIG. 6. $(1-\alpha)/\alpha$ as a function of conductance. Films at temperatures of 29·9°C, 39·4°C and 49·4°C were illuminated with a background d.c. light.

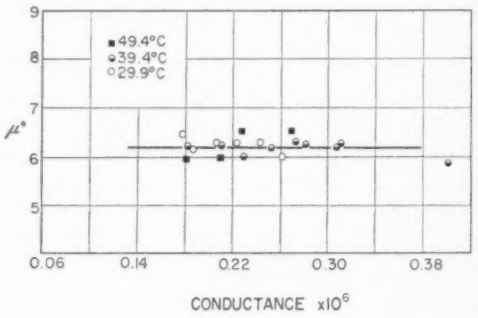


FIG. 7. Film mobility vs. film conductance for various levels of d.c. light illumination.

substantially linear relationship between $1/R$ and $(1-\alpha)/\alpha$ with the slopes tending to decrease with temperature. In Fig. 7, it is seen that μ^* in no way depends on the amount of illumination. This is still further proof that the barrier modulation hypothesis is incorrect for these films.[11]

The conductance $1/R = G_0$ of the film is plotted as a function of reciprocal temperature in Fig. 8. The resistance was quite stable and reproducible in the less photosensitive cells. As the photosensitivity increases, the tendency to drift similarly increases.

As found by SORROWS[8], the field effect and photo time constants were equal within experimental error. At high frequency the shape of the frequency vs. photoresponse curve deviated from the $1/\sqrt{(1+\omega^2\tau^2)}$ law. This can be attributed to the d.c. component of the pulsed light source which develops as $\omega\tau \gg 1$. This can best be seen diagramatically in Fig. 9. To understand the trapping and recombination processes occurring at the surface, it is necessary to establish the height of the surface barrier. The following evidence is a strong indication that the bands are flat, i.e. the surface potential is zero. One set of films prepared for this work was coated by the manufacturer with a plastic film. Another set made under identical conditions was left uncoated. The films were exposed to the laboratory atmosphere for a period of two years. The lifetime of the coated films was quite reproducible over this period of time whereas the lifetime of the uncoated films

changed. The field effect lifetime became significantly smaller than the photoeffect lifetime. This result implies that the photoeffect "sees" the trapping of carriers on all sides of the crystallite while the field effect "sees" only the traps opposite

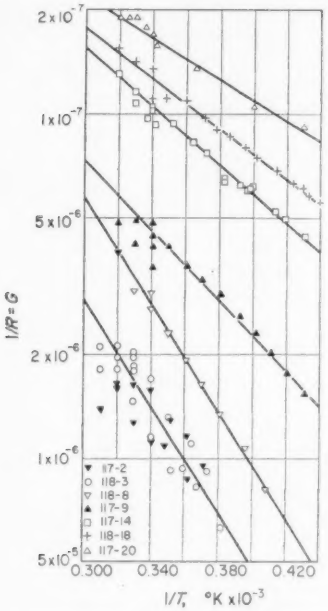


FIG. 8. Conductance vs. reciprocal absolute temperature.

the field plate. It is reasonable to assume that for sufficiently large values of the surface potential, the lifetime of surface carriers will drop by analogy with the surface recombination velocity. When the time constants are equal, we can then conclude that the bands at the surface are essentially flat.

A cursory study of ambients showed that the effect of water vapor on the resistance and the probability of capture ($1/\tau_{12}$) was quite pronounced. At sufficiently high water vapor pressures, some

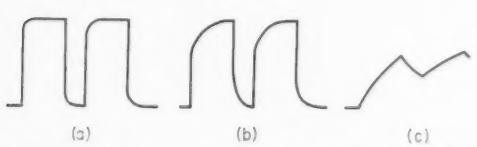


FIG. 9. Photoresponse as a function of time for various $\omega\tau$ values.

remarkable non-linear effects were observed along with an enormous rise in noise. No detailed studies were made of this phenomenon, although similar effects were observed when slow state studies had been carried out.[7] These water vapor effects were most pronounced when oxygen was present.

On two films of another series some ambient studies were carried out.[13] The results of mild changes of ambients (dry N_2 and O_2, wet N_2 and O_2) indicated that the field effect was sensitive to ambient change in PbS but not to the same extent that one would find for Ge or Si.

SINGLE TRAP MODEL OF PbS FILMS

Following SHOCKLEY and READ[10], one may write for the capture and emission rates

$$U(\text{capture}) = qN_{ts}^{e}\Delta pC_p/d \tag{9a}$$

$$U(\text{emission}) = qp_1C_p\Delta p_t \tag{9b}$$

where N_{ts}^{e} is the density of empty surface states, d is the sample thickness, C_p is the capture probability for holes times the thermal velocity, Δp and Δp_t are excess free and trapped holes respectively and p_1 is the number of holes if the Fermi level is at the energy of the trap.

$$\frac{U_{\text{capt}}}{\Delta p} = \frac{1}{\tau_{12}} = N_{ts}C_p \exp\left[-(\epsilon_t-\epsilon_f)/KT\right]/d \tag{10}$$

$$\frac{U_{\text{em}}}{\Delta p_t} = \frac{1}{\tau_{21}} = pC_p \exp\left[-(\epsilon_t-\epsilon_f)/KT\right] = p_1C_p \tag{11}$$

and $N_{ts} = (\alpha/1-\alpha)\,pd$ is the density of occupied surface levels.

DISCUSSION

In Table 2 various observed and derived quantities are tabulated for the films. There are a number of conclusions which can be drawn from the data in Table 2 and the other experimental data. We will treat these in turn.

Carrier concentration

Examination of the p vs. T curves, Fig. 10, indicates that the sensitive films had lower carrier densities and the statistics were essentially Maxwellian in behavior. As the sensitivity (and resistance) of the films dropped, the carrier concentration increased and the temperature dependence of the carrier concentration indicates some

Table 2. Energy parameters in electron volts

Cell	E_{μ^*}	E_p	E_{12}
117–20	—	0·039	0·124
118–18	0·038	0·039	0·124
117–14	0·054	0·041	0·156
117–9	0·061	0·047	0·219
118–8	0·057	0·094	0·241
118–3	0·054	0·113	0·241
117–2	0·054	0·113	0·241

degeneracy in the statistics. The least sensitive films show the most degenerate behavior. The activation energies vary from 0·10 eV to 0·05 eV.

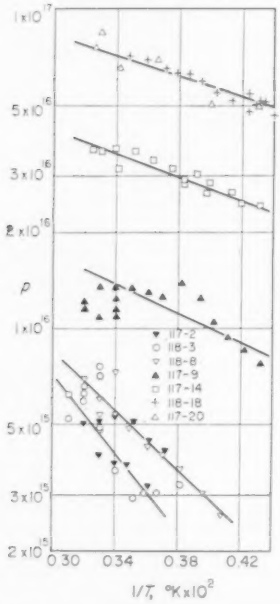


FIG. 10. Carrier concentration vs. reciprocal absolute temperature.

Trap densities

As the carrier concentration rises, the trap density also rises but not until near degenerate behavior is observed as shown in Fig. 11. In some cases there appears to be a temperature dependence of the number of occupied traps. However, the data on this point are open to the greatest amount of error.

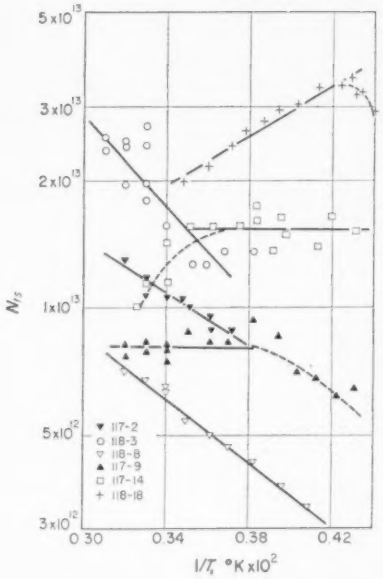


FIG. 11. Density of filled surface traps vs. reciprocal absolute temperature.

Mobilities

The mobilities of holes in the films as a function of temperature suggest an activation energy of 0·05 eV. The data are shown in Fig. 12. This gives rise to a serious problem. The deBroglie wavelength of the holes in PbS is of the order of 100 Å. Assuming an image potential model for the

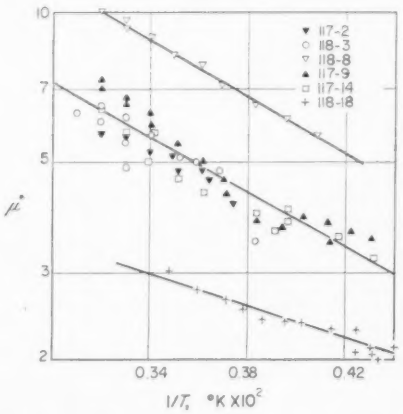


FIG. 12. Film mobility vs. reciprocal absolute temperature.

VOL. 14 1960

barrier between the crystallites, one finds that the barrier height is related to the thickness. From this one can show that the barriers are at most a few angstroms thick. The image potential approach is not valid at such small distances but assuming some kind of shorter range interaction, one obtains a value for the thickness of the barrier which is considerably less than the deBroglie wavelength of the holes. A detailed explanation of the mobility of thin films in general remains to be given.

Trapping cross sections

The capture cross section can be estimated roughly by using equation (10). The traps of film 117–9 are almost at the Fermi energy (hence the almost flat curve). Therefore, $\epsilon_t - \epsilon_f \approx 0$. Inserting the values of $1/\tau_{12}$, N_{ts} and d, one finds that

$$C_p = \sigma_p V \approx 1{\cdot}8 \times 10^{-14} \text{ cm}^3/\text{sec}$$

a similar estimate of C_p for 118–18 yields

$$C_p = \sigma_p V \approx 3 \times 10^{-15} \text{ cm}^3/\text{sec}$$

The velocities are of the order of 10^6–10^7 cm/sec which leads to a cross section, σ_p, of 10^{-21} cm^2. A quick examination of the temperature dependence of the various terms, i.e. N_{ts}, $\exp(\epsilon_t - \epsilon_f)/kT$ and d quickly established that most of the temperature dependence of τ_{12} and τ_{21} would be due to the capture cross section in this model.

The very small magnitude and large temperature dependence of the cross section is rather surprising. A similar situation occurs with nickel centers in Ge where very small cross sections have been observed. OKADA[12] has explained this by using a two-level system. Since the photoconductive studies on the lead salts strongly associate oxygen with the occurrence of photosensitivity, and since a complex molecule of oxygen, lead and sulfur might be expected to have more than one level in the gap, it is quite possible that a detailed analysis such as the one OKADA used might reveal the energy structure of the trapping center.

Minority carriers

While the field effect and photoconductivity in these films give abundant information on the majority carrier phenomena, the situation with respect to the minority carriers, i.e. electrons, is just the reverse. Virtually nothing is known about these carrriers at present. From measurements of the generation–recombination noise in PbS,[9] we know that recombination occurs at the same centers responsible for both photoconductivity and field effect trapping. Thus any minority carrier effects must be much slower than the majority carrier phenomena. Before a complete model of photoconductivity in PbS can be elucidated, more information will be needed here.

CONCLUSIONS

The equality of the field effect and photoconductive time constants over a wide range of temperatures and film characteristics confirms the results of SORROWS. On the basis of a simple Shockley–Read trapping model a number of parameters can be deduced such as trap densities and cross sections. While there is a great deal of consistency in the data, there are difficulties connected with understanding the mobility of the carriers, the role of the minority carriers and the apparent temperature dependence of the cross sections. However, thin films of PbS appear to be exceptionally well suited for surface studies of induced surface states.

REFERENCES

1. RITTNER E. S., *Photoconductivity Conference* (edited by BRECKENRIDGE, RUSSELL and HAHN) p. 215. Wiley, New York (1956).
2. SLATER J. C., *Phys. Rev.* **103,** 1631 (1956).
3. PETRITZ R. L., *Phys. Rev.* **104,** 1508 (1956).
4. BODE D. and LEVINSTEIN H., *Phys. Rev.* **96,** 259 (1954).
5. HUMPHREY J. N. and SCANLON W. W., *Phys. Rev.* **105,** 469 (1957).
6. PETRITZ R. L., LUMMIS F. L., SORROWS H. E. and WOODS J. F., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 229. University of Pennsylvania Press, Philadelphia (1957).
7. ZEMEL J., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 193. University of Pennsylvania Press, Philadelphia (1957).
8. SORROWS H., Ph.D. Thesis, Catholic University of America (1958).
9. PETRITZ R. L. and LUMMIS F., *Phys. Rev.* **105,** 502 (1957).
10. SHOCKLEY W. and READ W. T., *Phys. Rev.* **87,** 835 (1952).
11. WOODS J. F., *Phys. Rev.* **106,** 235 (1957).
12. OKADA J., *J. Phys. Soc. Japan* **12,** 741, 1338 (1957).
13. VARELA J. O. and ZEMEL J. N., *Bull. Amer. Phys. Soc.* ser. 2, **3,** 255(1958).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 150–154. Printed in Great Britain.

SURFACE SENSITIVITY OF RECTIFICATION

V. G. BHIDE

Physics Department, Institute of Science, Bombay, India

Abstract—The diffusion coefficient of sulphur into PbS, FeS_2 and other sulphide semiconductors has been studied as a function of temperature by employing the etching and light profile microscope techniques. The penetration of sulphur into an *n*-type crystal brings about various changes in the electrical parameters of the sample. It is possible to obtain by controlled treatment layers, with the requisite carrier concentration for the study of photovoltaic effect and transistor action. It has been possible to observe in PbS photovoltaic effect and transistor action with current gain equal to 0·83 and voltage gain of about 4.

With treatment of natural and artificial *n*-type crystals, a definite improvement in rectifying characteristics has been observed. The slope α of the semilog plot of such treated samples is nearly 40 V^{-1} at room temperature as expected by the theory. The slope α is usually found to depend on ϕ which in turn varies from point to point on the crystal surface and from layer to layer in the treated samples. This point-sensitivity as observed on the surfaces of a number of crystals is attributed to the crystal imperfections which terminate on the surface. Similar results have been observed employing a single crystal of germanium grown by Marconi Wireless and Co. This crystal surface (111) is seen to be covered with a mosaic structure, the figures being hexagonal in shape. These figures can be interpreted as a row or a lattice of dislocations as pointed out by Burgers.

The effect of screw dislocations on the electrical properties has been investigated using SiC crystals. An interesting correlation of the effect of dislocation on the rectifying characteristics has been indicated.

1. INTRODUCTION

It is well known that if a crystal of a binary compound such as PbS is heated in the vapour of one of the constituent elements, the gaseous atoms of that element penetrate into the crystal lattice by diffusion and occupy vacant lattice sites or deposit into interstitial positions. Reimann and Sullivan[1] have reported that by heating PbS in sulphur vapour, diffusion of sulphur into the crystal could be affected. Brebrick and Scanlon[2] have further shown that at a given temperature the number of sulphur atoms that diffuse into the crystal falls off exponentially towards the interior. It is easy to see that a controlled treatment of *n*-type crystals in the atmosphere of sulphur vapour could create a *p–n* junction in the body of the crystal.

The theories of rectification[3–7] at the semiconductor metal contact expect a slope α equal to e/kT or nearly 40 V^{-1} at room temperature in the semilog plot of *i–v* characteristics. Extensive analysis of germanium and silicon diodes by Yearian[18] and others give the value of α ranging from 5 to 35 V^{-1}. Although much work has been done on the diode and triode action in PbS[9–13] no attempt has so far been made to compare the observed characteristics with the theoretical model. It was of interest to study, therefore, the natural and treated samples especially because of the presence of the *p–n* junctions, both as regards rectification and transistor action.

2. EXPERIMENTAL

Natural and synthetic crystals which showed *n*-type conductivity were heated to about 500°C in an atmosphere of flowing H_2S for various lengths of time. The crystals were then quenched. These treated samples showed *p*-type conductivity. Similar conversion of *n*-type material to *p*-type and vice-versa was first reported by Eisenmann[14] and Hintenburger[15]. These workers were concerned with the bulk conversion and hence subjected their samples to severe sulphurization. We have on the other hand, treated the crystals for different durations leaving thereby a *p–n* junction in the

VOL.
14
1960

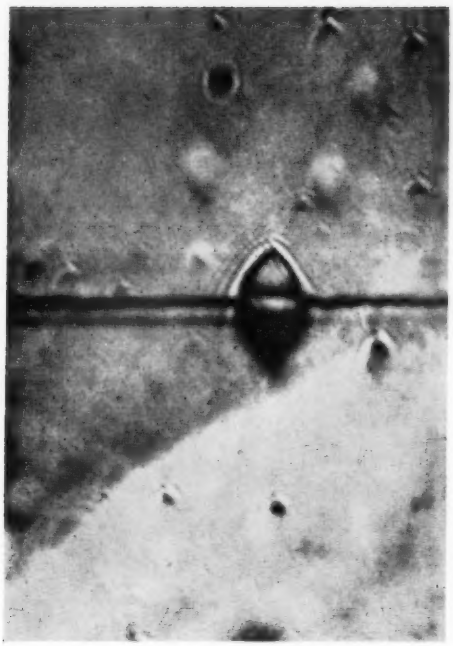


FIG. 1. Light profile over an etch pit. ×600.

[facing p. 150

VOL
14
196

body of the crystal. The depth of the p–n junction thus formed was determined by etching the crystal and testing the sign of rectification, thermal e.m.f. at different depths. The depth of etchpits were determined by light profile microscopy.[16] In this technique the field of view of the microscope is crossed by a dark profile line which is the image of an opaque line placed at the field iris and projected on the surface by an off-centre illumination. The profile gives the line contour of the specimen wherein changes in level are registered as lateral (Fig. 1) shifts in the profile. The depth of the p–n junction was found to be a linear function of time tending gradually towards saturation for a given rate of flow and crystal temperature. The points which made good rectifying contacts were investigated. Whiskers of different material made from wires of 1–3 mm and pointed by electrolytic etching were engaged to the crystal under optimum pressure.[17]

To study the transistor action, the treated crystals were cleaned and the base was ground for embedding in Wood metal to establish a non-rectifying base contact. To determine the extent of sulphur penetration, the diode characteristics were displayed on the CRO. A sharpened phosphor bronze whisker was adjusted by a micromanipulator to engage the crystal. Values of R_s were determined from the linear portion of the forward characteristic displayed on the CRO. The successive exposed layers formed by gently scraping the crystal with a sand paper were studied and the layer with appropriate carrier concentration was obtained for studying the transistor action.

Photoeffects were observed by illuminating the point contact with a fine spot of light from a 3-W bulb, the light being focussed with a microsope objective. The short-circuited photocurrent on illumination was measured by a galvanometer of current sensitivity of 12 div/μA. The heating effects at the contacts were eliminated by illuminating the point contact for a very short time.

To study i_c–v_c characteristics with varying d.c. emitter current i_e, reverse characteristics of the collector were observed on the CRO screen (Fig. 2). Care was taken to see that the emitter–collector resistance was high. This was ascertained from the i–v characteristics between them displayed on the CRO. The transistor action data was obtained by the use of electronic voltmeter and milliammeter of suitable ranges. From this data, the d.c. transistor characteristics were plotted.

3. RESULTS

The conductivity of natural n-type specimen was found to be of the order of 400 Ω^{-1} cm^{-1} which suggests that the crystals are sufficiently impure to inhibit carrier injection. This is further borne out by the large reverse currents and poor diode characteristics. In the absence of such carrier injection, the spreading resistance R_s and the barrier height ϕ were determined by extrapolating the linear portion of (i–Va) curves in the easy flow

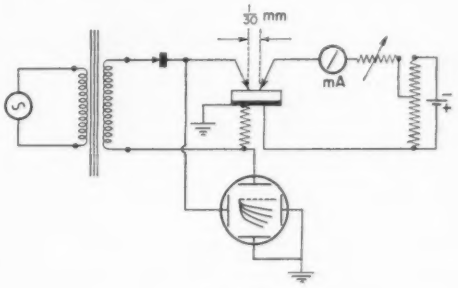


FIG. 2.

direction. The R_s thus determined and its corresponding values obtained following GIBSON[12] are in good agreement. It is interesting to note that ϕ decreases on partial sulphurization. It, however, increases again on complete sulphurization. The slope α of the semilog plot $[i-(Va-iR_s)]$ for n-type material was about ten, a quarter of the theoretically expected value. Unlike the n-type specimen and also Ge and Si diodes, the partially converted samples yielded the theoretical value of α with nearly all the whiskers studied. The slope of the linear portion of the reverse semilog curve $\beta e/kT$ (COURANT[6]) agrees fairly well with its calculated value from the forward characteristics (Table 1). It is also of some interest to note that the extrapolated zero voltage currents in the forward direction $(i_0)_F$ and the saturating currents in the reverse direction $(i_0)_b$ are of the same order. The treated crystals could stand higher breakdown voltage and they exhibited stabler characteristics. These results showed that the partially converted crystals approximated to the theoretical model.

Table 1

Crystal	Whisker	T(°K)	Slope	$S(T/e)$	$R_s(\Omega)$	$S(T/e)$ (V)	$(i_0)_f$ (A)	e/kT	B	$1-B$	$(i_a)_b$ (μA)
A_1P	Ni	306·4	35	0·91	161	0·18	2·1	1·8	0·045	0·95	—
A_1P	Ni	311·0	32·8	0·85	125	0·25	0·11	2·54	0·066	0·93	0·2
A_1P	Phos. Broz.	311·0	35·7	0·93	352	0·18	0·29	0·98	0·025	0·98	0·35
A_1P	W	311·0	37·0	0·96	147	0·19	0·31	0·95	0·025	0·98	0·45
A_1P	W	310·9	38·4	0·99	116	0·20	0·23	3·07	0·080	0·92	0·2
A_2P	Ni	313·8	32·0	0·84	549	0·22	0·50	2·88	0·075	0·93	0·50
A_2P	Ni	313·8	32·5	0·85	440	0·26	0·05	4·19	0·110	0·89	0·05
A_2P	W	313·7	37·7	0·98	448	0·19	0·11	1·99	0·052	0·95	0·20
A_2P	Phos. Broz.	313·4	38·9	1·02	1300	0·15	0·15	2·53	0·066	0·93	0·20
A_3P	Ni	313·2	41·1	1·07	175	0·12	0·30	0·57	0·015	0·99	1·0
A_4P	W	312·6	38·7	1·01	128	0·25	0·001	—	—	—	0·15
A_4P	W	313·0	40·0	1·04	680	0·16	0·16	—	—	—	0·30
A_4P	Phos Broz.	313·8	39·7	1·04	480	0·22	0·06	1·61	0·042	0·96	0·10
I N	Ni	306·0	9·9	0·25	53	0·52	12·0	3·85	0·098	0·90	—
II N	Ni	306·0	3·6	0·09	43	0·75	77·3	4·50	0·12	0·88	—

P = p-type; N = n-type.

YEARIAN[18] has explained the divergence between the observed and the theoretical values of α by suggesting that the contact covers a number of points with different values of ϕ. This postulate does not lead to unique determination of ϕ and its distribution within the contact area. The data presented here indicates that if ϕ is less than 0·25 V, the slope of (ln i–V) curves almost equals the theoretically expected value e/kT. Conversely, if ϕ exceeds this value as for n-type and fully converted crystal surfaces, the slope is much less than e/kT as is reported by several workers for Ge and Si diodes. A similar analysis of commercially available Ge diodes OA 70, OA 71, OA 74, OA 75 and freshly cleaved and treated samples of FeS_2 crystals has shown that α depends on ϕ and further if ϕ is less, then α approaches the theoretically expected value e/kT (Table 2). It is also of interest to note that the slope (ln i–V) for Ge and Si diodes changes rapidly from the theoretically expected value to much lower value after 0·2–0·3 V.[18] By partially treating the crystal, it was possible to produce a p–n junction in the body of the crystal. The sulphur atoms, during treatment, diffuse into the lattice and cause conversion of n-type material to p-type. Thus, by controlling sulphur penetration, it would appear possible to obtain layers of varying resistivity and select a suitable layer with requisite carrier concentration for the study of photovoltaic effect and transistor action.

On examining the section of the treated crystal with a probe, as explained earlier, the p-type characteristic could be detected down to a depth of 1/16 mm indicating the extent of penetration. The variation of R_s with depth from the uppermost treated surface is shown graphically in Fig. 3. As most of the contribution to R_s comes from the

Table 2

Crystal	Volts	R_s(V)	a	Temperature (°C)
FeS_2				
1 W	0·24	403	30·0	28·0
II W	1·15	596	9·0–3·2	27·4
III W	0·79	563	4·0	27·2
V W	0·40	351	19·0	27·2
VI W	0·32	688	30·0	28·4
Germanium				
OA 70	0·22	113·5	29·0	27·2
OA 71	0·23	173·0	27·0	27·2
OA 74	0·21	126·0	33·0	26·8
OA 74	0·17	134·0	42·0	27·2

VOL. 14 1960

portion of the semiconductor close to the contact, variation of R_s can be taken as an indication of the relative magnitudes of the local resistivity on the assumption that the contact area is the same in every case. It would thus appear that the layers with higher resistivity, i.e. with smaller carrier concentration are situated below the outermost treated surface. This is in accordance with the observation of BREBRICK and SCANLON[2] that the

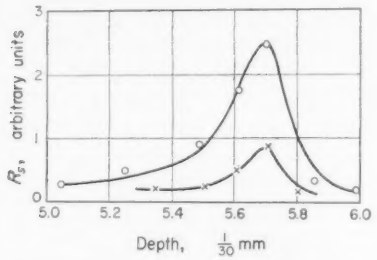


FIG. 3. Variation of R_s with depth.

number of sulphur atoms which diffuse into the crystal decreases exponentially with depth below the exposed surface and is dependent on temperature, duration of treatment and pressure of vapour. Addition of sulphur to *n*-type PbS creates acceptors as is seen from the reversal of the type of conductivity. During heating, the outermost surface is constantly exposed to sulphur vapour. It is thus over compensated and the conductivity is high due to the large concentration of holes. Towards the interior, the diffusion of sulphur decreases exponentially reducing the degree of compensation. The observed gradual increase in R_s towards the interior can thus be explained as due to reduced over-compensation and consequent smaller hole concentration. A layer where more or less perfect compensation has taken place is indicated by the highest values of R_s. The fall of R_s below such a layer is due to the contact passing over from *p*-type to *n*-type region, where the conductivity is high and is equal to that of the untreated specimen. The absence of photovoltaic effects and transistor action on H_2S treated galena surface after the heat treatment may be due to large hole concentration caused by over compensation of layers. The upper surface gave good *p*-type rectifying characteristics but did not show either photovoltaic effect or transistor action. On removing the surface layers by gently scraping or by etching the crystal the short circuited photo-currents could be detected. The maximum value of the observed photo-current was 1·66 μA. With a bias of about 2–3 V in the reverse direction, the photocurrent increased by about 12 per cent indicating the presence of carrier injection (these treated samples also showed the transistor action). The usual four transistor characteristics are shown in Fig. 4. The calculated parameters

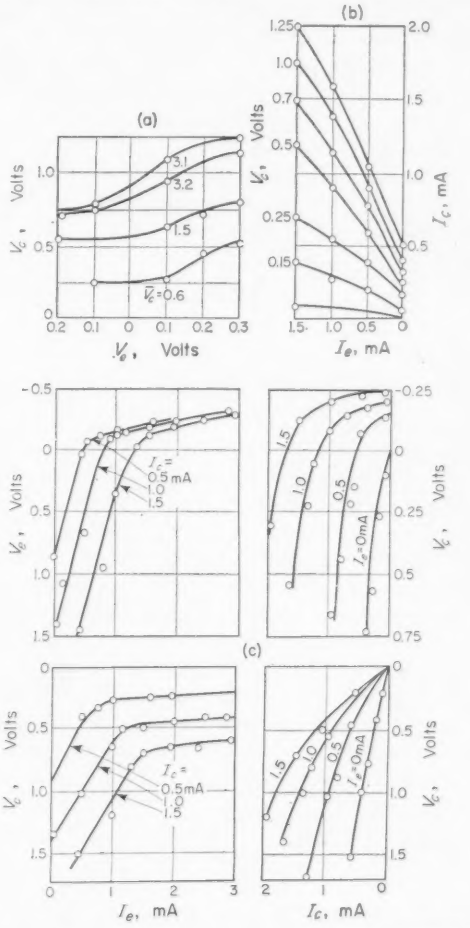


FIG. 4. (a) Voltage gain curves ($R_L = 1000\ \Omega$); (b) current gain curves; (c) transistor characteristics.

are $R_{11} = 106\ \Omega$; $R_{22} = 900\ \Omega$; $R_{12} = 125\ \Omega$; $R_{21} = 750\ \Omega$ and $\alpha = 0{\cdot}83$. It is to be noted that the values of the transistor parameters are an order smaller than the corresponding values reported

for good natural p-type crystals, though the magnitude of current gain obtained in the present investigation compares favourably with the values (0·33 to 0·9) reported by BANBURY *et al.*[19]

During all this investigation it was found that the crystal surface exhibited spot sensitivity in respect of the values of α, R_2 and ϕ. The observations could be explained by taking into consideration the effect produced by trapped immobile electrons in the surface states together with the usual barrier at metal–semiconductor contact. In the region where the density of trapped electrons is high, strong repulsive field for approaching electrons can be assumed to be present and as a consequence, points situated in this region have a higher value of ϕ. Assuming a certain transmission coefficient for the electrons that cross the barrier into the metal at the contact, the diffusion equation of MOTT could be written as

$$j = \beta nevF - De\, dn/dx$$

which gives

$$j = \beta nevF \exp -\beta e\phi/kT\,(\exp \beta eV/kT) \text{ and}$$

$$\alpha = \beta e/kT$$

On comparing the modified diffusion equation with the experimental results, β has been found to vary between 1 and 0·2. This point sensitivity is, at least in part, attributable to the crystal imperfections which terminate on the crystal surfaces. This led to an investigation of the effects of dislocations on rectifying properties of various crystals. A single crystal of germanium grown by Marconi Wireless and Co. was examined by phase contrast microscopy and its surface topography was studied with multiple beam interference techniques. The entire surface (111) is seen to be covered with a mosaic structure the figures being hexagonal in shape varying in size and also observed to be arranged in a row.

The fringes of equal chromatic order have shown that the hexagonal figures are indeed depressions and can be interpreted as a row or lattice of dislocations as pointed out by Burgers. These germanium surfaces yielded a fairly low value of α indicating the existence of a large number of surface states. Indeed Shockley has shown that an edge dislocation in germanium would mean one dangling bond for every germanium atom in the extra plane. These bonds are responsible for acceptor type behaviour. When dislocations are near screw orientation, it is probably a good approximation to consider the dislocation acceptors as uniformly distributed. As these sites accept electrons, the dislocation line gets charged and space charge region forms round the dislocations. The effect of such screw dislocations on the rectifying properties of crystals has been investigated using SiC crystals. The crystal surfaces exhibit a variety of spiral markings originating essentially from screw dislocations. It is invariably observed that the value of α is lower at the point of emergence of the dislocations than elsewhere. Although a definite relation between the strength of the dislocation (determined by multiple beam interference techniques) and the corresponding lowering of α cannot at this stage be given, yet it appears that the higher the dislocation strength the lower is the value of α.

REFERENCES

1. REIMANN A. L. and SULLIVAN J. V., *Proc. Phys. Soc., Lond.* **B65,** 480 (1952).
2. BREBRICK R. F. and SCANLON W. W., *Phys. Rev.* **96,** 598 (1954).
3. MOTT N. F., *Proc. Roy. Soc.* **A117,** 27 (1939).
4. SCHOTTKY W., *Z. Phys.* **118,** 539 (1941).
5. BETH H. A., *N.D.R.C.*, Div.14, Rep. No. 43 (1942).
6. COURANT E. D., *Phys. Rev.* **69,** 684 (1946).
7. BANBURY P. C., *Proc. Phys. Soc., Lond.* **B66,** 833 (1953).
8. YEARIAN H. J., *J. Appl. Phys.* **21,** 214, 283 (1950).
9. HENISCH H. K. and GRANVILLE J. W., *Semiconducting Materials*, p. 87, Butterworths, London (1951).
10. BANBURY P. C. and HENISCH H. K., *Proc. Phys. Soc., Lond.* **B63,** 540 (1950).
11. GEBBIE H. A., BANBURY P. C. and HOGARTH C. A., *Proc. Phys. Soc., Lond.* **B63,** 371 (1950).
12. GIBSON A. F., *Proc. Phys. Soc., Lond.* **B65,** 196 (1952).
13. SCANLON W. W., *Phys. Rev.* **99,** 620 (1955).
14. EISENMANN L., *Ann. Phys. Lpz.* **38,** 121 (1940).
15. HINLINBURGER H., *Z. Phys.* **119,** 1 (1942).
16. TOLANSKY S., *Z. Electrochem.* **56,** 263 (1952).
17. DAS J. N. and BHIDE V. G., *Curr. Sci.* **23,** 185 (1954).
18. TORREY H. C. and WHITMER C. A., *Crystal Rectifiers*. McGraw-Hill, New York (1948).
19. BANBURY P. C., GEBBIE H. A. and HOGARTH C. A., *Proc. Phys. Soc., Lond.* **B63,** 371 (1950).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 155–161. Printed in Great Britain.

THE CHARGE DISTRIBUTION AT THE ZINC OXIDE–ELECTROLYTE INTERFACE

J. F. DEWALD

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

VOL. 14 1960

Abstract—Capacitance measurements made on single crystal zinc oxide electrodes in contact with aqueous electrolytes are reported. Over a wide range of bias and bulk donor density, the results are in almost quantitative accord with predictions of the simple Poisson–Boltzmann (Poisson–Fermi in the degenerate case) equation. This is shown to imply the complete absence of surface-state effects in this system. A very sharp discontinuity in the flat-band potential is observed at bulk electron densities in the range from $0{\cdot}6\times10^{18}$ to 2×10^{18} donors/cm^3. This and other effects, arising under varying surface treatments, are discussed in some detail.

1. INTRODUCTION

THE modern era in the study of the semiconductor electrolyte interface dates back to BRATTAIN and GARRETT's work[1] on the germanium electrode. Practically all of the studies since that time have been on germanium. This work is summarized in two recent review articles.[2, 3] Although the qualitative understanding of the germanium electrode is fairly well advanced complications arise because of the chemical reactivity of the material. This leads, in aqueous solutions at least, to the presence of oxide films and in addition precludes study of the electrode under equilibrium conditions.*

Zinc oxide was chosen for the present study because of its complete thermodynamic stability in contact with aqueous solutions of intermediate pH, and because next to germanium and silicon, its bulk properties (band structure, mobilities, impurity ionization energies, etc.) are the most thoroughly understood of all semiconductors. We will be concerned in the present paper with the distributions of charge and potential at the zinc oxide–electrolyte interface.

Figure 1 shows a diagram of a typical semiconductor electrode in contact with an electrolyte. For specificity we take an n-type semiconductor under conditions of anodic bias (bands bending up). The conventional energy-band diagram for the system is also shown in Fig. 1. In general, the Galvani (or inner) potential difference (ψ_0) may distribute itself over three regions of the interface; (1) the space-charge layer in the semiconductor, (2) the one or two atom-diameter region of the Helmholtz layer, and (3) the diffuse Gouy layer in the electrolyte. Since the Gouy layer has been much investigated in studies of metal electrodes and is no different at a semiconductor electrode, we will restrict ourselves in both our experimental and theoretical treatments to conditions (namely high ionic concentration in the electrolyte) under which the potential drop across the Gouy layer is negligible. The potential drop across the space-charge layer in the semiconductor, ψ_s, is called the surface potential by physicists.

The nature of the charge and potential distributions has been determined by measuring the differential capacitance of the zinc oxide electrode as a function of bias. Using the nomenclature of Fig. 1, the capacitance (C) of the electrode may be written

$$C = -\frac{dQ}{d\psi_0} = -\left\{\frac{dQ_{\text{s.c.}}}{d\psi_s} + \frac{dQ_{\text{s.s.}}}{d\psi_s}\right\}\frac{d\psi_s}{d\psi_0} \quad (1)$$

where Q, $Q_{\text{s.c.}}$ and $Q_{\text{s.s.}}$ are, respectively, the total charge on the semiconductor, the charge in the space-charge layer and the charge in "surface states". This corresponds to having a large ($\sim 30\,\mu f/\text{cm}^2$)[5] capacitance C_H in series with two

* See, for example, the work of BOHNENKAMP and ENGEL[4].

others ($C_{s.c.}$ and $C_{s.s.}$) in parallel and the total capacitance is given by

$$\frac{1}{C} = \frac{1}{C_H} + \frac{1}{C_{s.c.} + C_{s.s.}} \tag{2}$$

where $C_{s.c.}$ and $C_{s.s.}$, in effect capacitances of the space-charge layer and of the surface states, are respectively $dQ_{s.c.}/d\psi_s$ and $dQ_{s.s.}/d\psi_s$. Now $C_{s.c.}$ may be computed from first principles (see below) and C_H estimated from results with metal electrodes so that a measurement of C gives an estimate of $C_{s.s.}$.

The theory of the space-charge layer at a semiconductor surface has been given by a number of workers.[6-8] For donors completely ionized in the bulk of the crystal $C_{s.c.}$ may be written

$$C_{s.c.} = \left(\frac{q^2 N_D \kappa\kappa_0}{2kT}\right)^{1/2} \frac{|e^Y - 1|}{\{e^Y - Y - 1\}^{1/2}} \tag{3}$$

A number of special cases of equation (3) are of special interest. When Y is strongly negative, i.e. when the bands are bent up quite strongly, we have

$$C_{s.c.} = \left(\frac{q^2 N_D \kappa\kappa_0}{2kT}\right)^{1/2} \{-1 - Y\}^{-1/2} \tag{4}$$

In the limit of large negative Y this is just the Mott–Schottky approximation,[6, 7] and a plot of $1/C_{s.c.}^2$ vs. Y should give a straight line, whose slope is a measure of the total donor density at the surface.

At the other extreme of potential, when the bands bend down and Y is large and positive, equation (3) may be approximated in the form

$$C_{s.c.} = \left(\frac{q^2 N_D \kappa\kappa_0}{2kT}\right)^{1/2} (e^Y)^{1/2} \tag{5}$$

As can be seen the capacitance increases, apparently without limit, as the bands bend down and

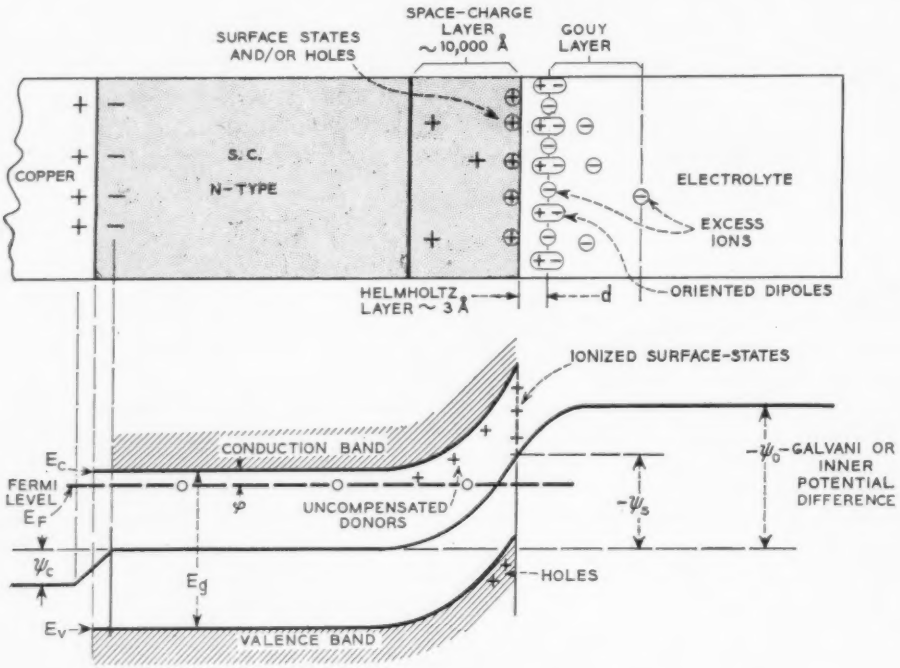


FIG. 1. Schematic structure and energy level diagram of a typical n-type semiconductor–electrolyte interface.

the electrons tend to crowd up closer and closer to the surface. Eventually, however, the surface layers become degenerate and then the Poisson–Boltzmann approximation breaks down. This region is considered in greater detail below.

2. EXPERIMENTAL PROCEDURES

(a) *Sample preparation*

The crystals used in this study were grown in these laboratories by D. G. THOMAS and R. T. LYNCH using the vapor phase reaction between zinc and oxygen at about 1200°C as originally reported by SCHAROWSKY[9]. The resulting crystals were in the form of hexagonal needles, about 0·1–0·3 mm in "diameter" with the primary faces being {11·0}. Conductivities of the "as-grown" crystals were in the range from about 0·01 to 3·0 Ω^{-1} cm^{-1}. The crystals were selected for their uniformity of composition both lengthwise and radial. The lengthwise uniformity was checked by conductivity measurement, while the radial uniformity was checked by measurement of both conductivity and electrode capacitance between successive etches in 85 per cent H_3PO_4. Of a number of etching solutions employed this was found to be the most satisfactory by virtue of its uniform attack and its convenient rate (~ 1 μ/min at room temperature).

In the lowest region of conductivity the crystals were used in the "as-grown" condition. For conductivities in excess of about 1 Ω^{-1} cm^{-1}, the crystals were "doped" by high temperature diffusion of indium into the crystal. By far the largest numbers of crystals were doped with indium in a manner similar to that described by THOMAS[10]. This procedure was rather tedious for conductivities below about 10 Ω^{-1} cm^{-1} because of the extended equilibration times which would have been required. In this conductivity range the crystals were partially equilibrated with an excess of $In(NO_3)_3$ until the average conductivity had the desired value. The excess indium was then removed by etching in H_3PO_4, and the crystal replaced in the oven, usually at 1100°C, for times calculated (from THOMAS' diffusion data[10]) to give a maximum radial variation of 20 per cent or less.

Mechanically strong contacts were made to the crystals by electroplating a thin layer of indium on the tip of the crystal, then electroplating a second layer of copper on top of the indium, and finally soldering a copper wire to the copper plate. A fairly large glob of solder was used which completely surrounded the tip of the crystal. This clamps the crystal so tightly that on attempting to break the contact the crystal breaks instead. The resistance of the contact was negligible compared to the crystal resistance over the entire range of conductivity from about 0·005 to 150 Ω^{-1} cm^{-1}.

The solder joint and copper wire were masked by slipping the copper wire through a pyrex glass tube and then sealing with Apiezon *W* wax. The efficacy of the masking procedure was checked by measuring the d.c. current flow under anodic bias up to about 10 V. If any of the copper, solder, or indium is exposed to the solution, catastrophic currents flow under such a bias. On the other hand, a properly masked crystal passes currents only of the order of 10^{-9} A/cm^2 or less, almost completely independent of the crystal resistivity.*

(b) *Electrochemical cell and capacitance bridge*

The electrochemical cell employed was of fairly standard design. It was made of quartz and contained a large area working electrode of platinized platinum, a normal calomel electrode separated from the main solution by a 1N KCl salt bridge, and a second platinum electrode for pre-electrolysis. A nitrogen atmosphere was maintained over the solution at all times. The solution was a borate-buffered 1 N solution of recrystallized KCl (pH $\sim$ 8·5). Freshly ground, spectroscopically pure, zinc oxide with an estimated area of about 10,000 cm^2 was added to the solution as a "getter" for any surface-active impurities. Periodic additions of more zinc oxide were made.

Impedance measurements were made using a General Radio No. 716-*C* bridge with an auxiliary decade capacitance box in parallel with the variable air capacitor of the bridge. A schematic diagram of the circuit is shown in Fig. 2. C_x represents the capacitance of the zinc oxide electrode which was always less than about 0·3 μF. The platinum working electrode had an apparent area of about 10 cm^2 so that its capacitance was more than 10^3 times C_x and thus made a negligible contribution to the total capacitance.

* A study of the electrode kinetics at the zinc oxide electrode has been reported,[11] and details will be published shortly.

The electrode bias applied to the crystal was varied potentiometrically and measured, with respect to the calomel electrode, using a Leeds and Northrup vacuum tube voltmeter with an impedance of 1600 MΩ, and a sensitivity of about ±1 mV. The voltmeter was disconnected during

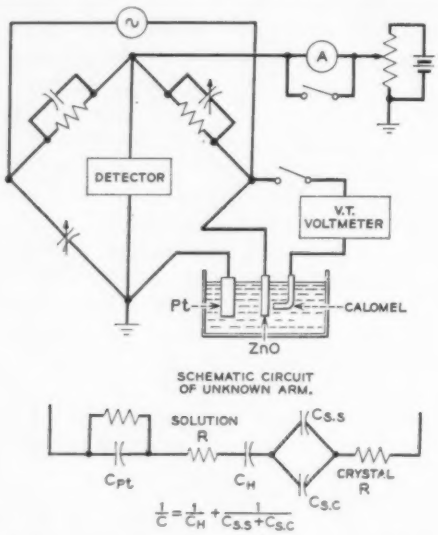


FIG. 2. Schematic diagram of the capacitance bridge and the equivalent circuit of the unknown arm.

capacitance measurement. The input signal to the bridge was kept below 10 mV. Measurements were made in the frequency range from 50 c/s to 100 kc/s. The bridge was extensively calibrated at the start of our experiments, and then checked periodically during the course of the measurements.

3. EXPERIMENTAL RESULTS

Exhaustion region

The results which have been obtained for the capacitance of the zinc oxide electrode in the exhaustion region, i.e. under anodic bias, are remarkable for their simplicity. The results for two crystals of intermediate conductivity are shown in Fig. 3. Here, taking our cue from equation (4), we plot the reciprocal of the capacitance squared against the electrode potential measured with respect to the calomel electrode. As can be seen the resulting plots are very nicely linear for the two crystals. This was the case for nearly every crystal studied; in only two crystals out of more than 50 was a curvature greater than 2 per cent observed in the region of positive potential (on the calomel scale) and most crystals showed a variation less than 0·5 per cent.

This behavior is precisely that predicted for the Mott–Schottky space-charge capacitance

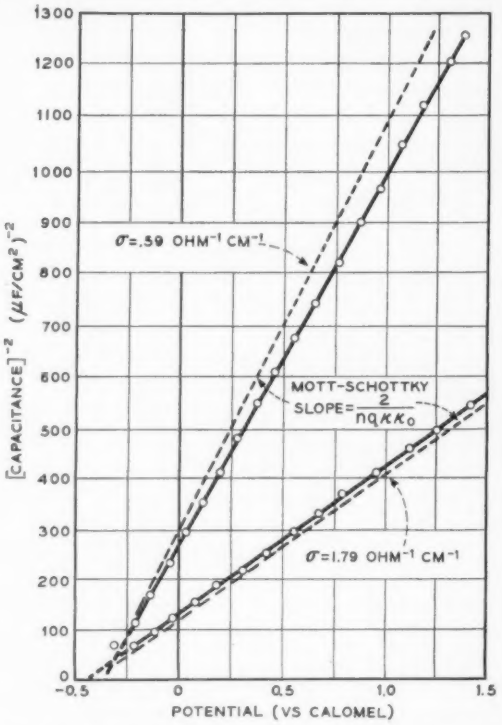


FIG. 3. Mott–Schottky plots for two crystals under exhaustion conditions. The dotted lines represent the theoretical slopes.

(equation 4) and implies that the surface-state capacitance is very much smaller than the space-charge capacitance. Note in this connection that a capacitance vs. bias curve for any *one* crystal can *always* be fit by a completely unrestricted distribution of surface states. Such an interpretation, while violating the law of simplicity, could not be denied. However, the same distribution should presumably apply for all crystals, independent of the impurity density, and one cannot devise a single surface-state distribution which will simultaneously satisfy the results on crystals of varying impurity density.

VOL. 14 1960

The magnitude of the slope of the $1/C^2$ vs. V plots confirms the simplicity of the results in the exhaustion region. Equation (4) says that, in the absence of surface-states, and if the donors are completely ionized (i.e. $N_D = n$) the slope should be given by

$$\text{Slope} = \frac{2}{nq\kappa\kappa_0} \qquad (6)$$

Since the dielectric constant of ZnO is known[12] ($\kappa = 8{\cdot}5$), the only unknown quantity is the electron density n which can be obtained from conductivity and Hall effect measurements. For the two crystals in Fig. 3 the agreement between the Mott–Schottky slope and the experimental slope is within experimental error,* as shown by the dotted lines whose slopes are calculated from equation (6). In very good approximation the data are frequency independent, the maximum variation being of the order of 1 per cent in the range from 50 to 100,000 c/s.

The conclusion that the surface-state capacitance is negligible compared to the space-charge capacitance is confirmed in all of our experiments. This is shown in Fig. 4, where we plot the slope of the $1/C^2$ vs. V curves as a function of electron density. The open circles represent "as-grown" crystals while the solid circles are for indium doped samples. The straight line in Fig. 4 is an absolute prediction with no adjustable parameters. As can be seen agreement of the data with this simple theory is fairly good over the entire range studied. It can be improved somewhat by realizing that at high density the indium donors are not completely ionized. HUTSON's preliminary measurements[13] on indium donors indicate an ionization energy of about 0·075 eV at electron densities around 10^{17} or less. If we assume that this value applies over the entire range of concentration with (see Ref. 12) $D = 2$ and $[m(N)]/m = 0{\cdot}5$ we obtain the curve marked "$E_D = 0{\cdot}075$ eV" Since the donor ionization energy is expected to decrease with increasing donor density we expect our experimental points to lie beween the two curves in Fig. 4, as in fact they do.

The effective absence of surface states at the zinc oxide–electrolyte interface implies that the electric field arising from free charge on the semiconductor cannot give rise to any significant potential drop across the Helmholtz layer. Sizable changes of the surface dipole, on etching a crystal or on long term equilibration, are indicated, however, by our measurements of the "flat-band potential," i.e. the electrode potential at which the bands are flat. From equation (4) we see that the extrapolation of linear portion of the $1/C^2$ vs. V plot to the voltage axis gives the potential at which the bands bend up by kT/q V, from which one readily obtains the flat-band potential.

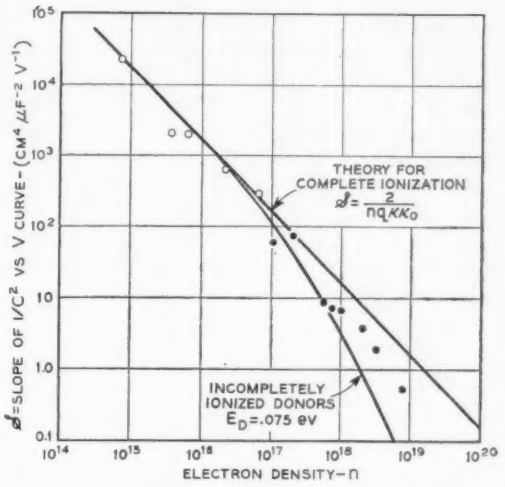


FIG. 4. The Mott–Schottky slope as a function of the bulk electron density. The straight line is taken from equation (6), and represents a no-parameter fit to the data. The curved line shows the correction for incomplete donor dissociation with $E_D = 0{\cdot}075$ eV. ○ "as-grown" crystals, ● indium doped crystals.

Figure 5 shows the experimental variation of the flat-band potential with the bulk electron density for three different surface treatments. The curves marked "H_3PO_4" and "KOH" were taken within a few minutes after a brief (2–10 sec) etch in, respectively, H_3PO_4 (85 per cent) and KOH (3 M). The curve marked "Long Stand" was obtained with crystals which had stood in the buffered KCl electrolyte (pH ~ 8·5) for 10 hr or more.

Consider the data at low concentration first. There are two possible sources of a variable flat-band potential, (1) a variable surface dipole and (2) a variable contact potential difference at the copper–zinc oxide interface due to the variation

* The major errors arise in the determination of surface area and in the determination that the crystals are in fact uniformly doped.

with electron density of the bulk Fermi level. The latter variation goes as $(kT/q)\ln n$ for a nondegenerate semiconductor and the straight lines at low electron density are drawn with slope to correspond to this variation (59 mV/decade in n). As can be seen the change in contact potential accounts for most, if not all, of the observed variation with crystal doping at levels below about 10^{18} ions/cm^3. This is not particularly surprising since the impurity density at such levels is quite low in normal chemical parlance (99·99 mol.% ZnO) and one would not expect the chemistry of a proton going on to the surface on etching in acid and coming off when one etches in base. Since the proton gets much closer to the "surface" than does the compensating negative ion, the surface dipole will be more negative after a basic etch than after an acid etch.

Consider now the data in Fig. 5 at high doping levels. A remarkably sharp "bump" amounting almost to a discontinuity, in the flat-band potential is observed centered at a bulk electron density of about 10^{18}, with the flat-band potential getting more positive by about 130 mV on going from an

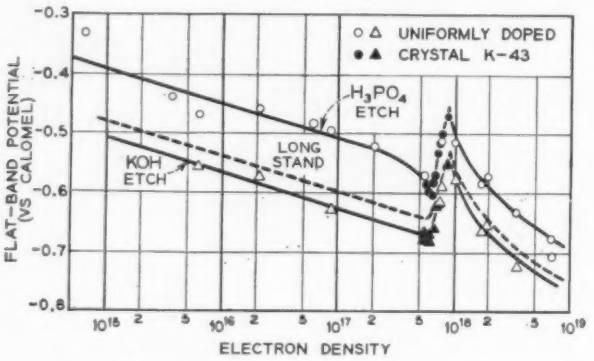


FIG. 5. The variation of the flat-band potential with bulk electron density for three different surface treatments.

the surface to be strongly impurity sensitive at these low levels.

A more physically interesting result is obtained from the variation of the flat-band potential with the nature of the surface treatment. As can be seen from Fig. 5 the flat-band potential is shifted downwards by about 130 mV when one etches in KOH instead of H_3PO_4, corresponding to the addition of a surface dipole with the positive end towards the electrolyte (or removal of a dipole of opposite polarity). This effect is quite reproducible, on a given crystal, and quite rapid; a 2-sec etch in either H_3PO_4 or KOH (corresponding to the removal of about 300 Å and 30 Å in the two cases) brings the flat-band potential to the values shown, completely independent of the past surface treatment.

The 130 mV change in flat-band potential observed on going from a KOH etched surface to an H_3PO_4 etched surface may be qualitatively understood as a heterogeneous acid base equilibrium,

electron density of $0{\cdot}6\times10^{18}$ to $0{\cdot}9\times10^{18}$. Within experimental error the contour of the bump is independent of surface treatment. The "width" of the bump is so small that in our early experiments, which had only one crystal in the range from $0{\cdot}6\times10^{18}$, we completely overlooked it and took the low density behavior to be valid over the entire concentration range. We had then just a single anomalous point which we rationalized as arising from some unknown anomaly in chemical or thermal history of the crystal. This rationalization became quite tenuous as further crystals were studied in this concentration range and was completely eliminated by the data shown as shaded points in Fig. 5. To obtain these data we took a crystal which, from capacitance measurements, had total donor density less than 10^{16} donors/cm^3 and by a relatively short heat treatment with indium produced a nonuniform donor distribution, the outside of the crystal having a donor density

near 10^{18}, while the inside was at a level around 10^{17} donors/cm^3. Capacitance and conductivity measurements were then made as the crystal was successively etched. In this way we obtained a plot of flat-band potential vs. surface donor density with a minimum of ambiguity as regards uncontrolled variables of thermal and chemical nature. As can be seen from Fig. 5, these data are nicely consistent with the data obtained on uniformly doped crystals and demonstrate that the "anomalous" variation of flat-band potential with doping is a real effect.

The origin of the bump in flat-band potential is at present not known. Taken at face value the data say that the bulk Fermi level *decreases* as one *increases* the total electron density. While such a behavior is completely impossible for any classical system it might just possibly arise as a result of an impurity band phenomenon. If the energy of the impurity band decreased sufficiently rapidly with increasing doping it might conceivably draw electrons from the conduction band even though the total electron density increased. For want of a better explanation we can only present this as a preliminary hypothesis and wait for future experiments and theoretical interpretation to confirm it.

VOL. 14 1960

REFERENCES

1. Brattain W. H. and Garrett C. G. B., *Bell Syst. Tech. J.* **34,** 129 (1955).
2. Dewald J. F., *Semiconductor Chemistry* (edited by Hannay N. B.) Monogr. Ser. Amer. Chem. Soc., Reinhold, New York (1959).
3. Green M., *Modern Aspects of Electrochemistry* (edited by Bockris J. O'M.). Butterworths, London (in press).
4. Bohnenkamp K. and Engel H. J., *Z. Elektrochem.* **61,** 1184 (1958).
5. Parsons R., *Modern Aspects of Electrochemistry* (edited by Bockris J. O'M.) Butterworths, London (1954).
6. Schottky W., *Z. Phys.* **113,** 367 (1939); **118,** 539 (1942).
7. Mott N. F., *Proc. Roy. Soc.* A**171,** 27 (1939).
8. Garrett C. G. B. and Brattain W. H., *Phys. Rev.* **99,** 376 (1955).
9. Scharowsky E., *Z. Phys.* **135,** 318 (1953).
10. Thomas D. G., *J. Phys. Chem. Solids* **9,** 31 (1958).
11. Dewald J. F., *Abst. N.Y. Meetg. Amer. Electrochem. Soc.* (1958).
12. Hutson A. R., *Phys. Rev.* **108,** 222 (1957).
13. Hutson A. R., personal communication.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 162–168. Printed in Great Britain.

STRUCTURE AND ADSORPTION CHARACTERISTICS OF (111) AND (11$\bar{1}$) SURFACES OF InSb CLEANED BY ION BOMBARDMENT AND ANNEALING*

D. HANEMAN

Barus Research Laboratory of Physics, Brown University, Providence, Rhode Island

Abstract—Low energy electron diffraction studies have been made of five (111) and (11$\bar{1}$) surfaces of InSb which were cleaned by argon ion bombardment and annealing. From the diffraction patterns obtained in both (110) and (112) azimuths, a double-spaced structure of surface atoms appears to be present on (111) and (11$\bar{1}$) surfaces. Exposure to oxygen resulted in a gradual decrease in intensity of all beams, indicating that an amorphous layer of oxygen was formed. The sticking coefficient for oxygen on the cleaned surfaces has a maximum value of 10^{-5}, and appears to be smiliar on both (111) and (11$\bar{1}$) surfaces. The cleaned surface pattern could be regenerated by heating the crystal at about 330°C for approximately 1 hr.

1. INTRODUCTION

THEORETICAL considerations of surfaces have usually been based on a simple model whereby the surface is formed by the bulk structure ending on undistorted crystal planes. Recent structure determinations[1] of surfaces cleaned by the ion bombardment and annealing method[2] have however produced evidence that the surface atoms may be rearranged from bulk positions, especially for the diamond-structure semiconductors Ge and Si. We report here an extension of such measurements to InSb, a compound diamond-structure semiconductor of comparatively low melting point, 523°C. The results lend emphasis to the need to consider rearranged surface structures.

Because of the compound structure and low melting point of InSb, one feature of interest is whether the ion bombardment and annealing method can be used on this material to produce sharp low-energy electron-diffraction patterns. The breaking method of producing clean surfaces is only suitable for (110) surfaces of InSb,[3] whereas interest here was particularly in (111) surfaces. When two such surfaces are created by a cut parallel to a (111) plane they show different properties under chemical etching.[4] Hence there may be an asymmetry between such faces when atomically clean. An additional feature of interest is the gas adsorption characteristics of cleaned surfaces of InSb.

2. SURFACE PREPARATION

In all, five separate crystal surfaces were examined. The specimens were cut from a single crystal of InSb† of high purity (*p*-type, 7×10^{15} impurity atoms/cm^3) and the faces were oriented to within $\frac{1}{4}$ degree of (111) planes by an X-ray method. They were then ground and polished, and etched for 5 sec in a mixture of 2 parts 70 per cent HNO_3, 1 part 50 per cent HF and 1 part 98 per cent glacial acetic acid, diluted with distilled water. This was followed by a (111)-face preferential etch of equal parts of 50 per cent HF and 30 per cent H_2O_2 for 5 sec. The crystals were then rinsed in distilled water, and degreased in an acetone extractor prior to insertion in the experimental tube.

3. CRYSTAL MOUNTING

The crystals were studied in an experimental low-energy electron-diffraction tube‡ of a type

* This work was supported by a contract with the Air Force Cambridge Research Center, Air Research and Development Command.

† Kindly supplied by Dr. H. C. GATOS, Lincoln Laboratory, Lexington, Mass.

‡ Assembled by Dr. R. E. SCHLIER.

previously described by FARNSWORTH *et al.*[5] but with a different crystal mount. Because of the low m.p. of InSb, its temperature during annealing was monitored by a chromel–alumel thermocouple in permanent contact with the back of the crystal, as shown in Fig. 1. The crystal was held against

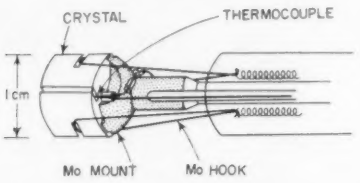


FIG. 1. Diagram of InSb crystal and its mounting.

small projections on the front of a molybdenum block by spring-loaded molybdenum hooks bearing on slots cut in the crystal sides. Because of the projections, the area of thermal contact between the crystal and block was small, permitting closer temperature control by heating the crystal directly (with electron bombardment) rather than by conduction from a heated mount.

4. VACUUM PROCEDURES

A diagram of the vacuum and gas handling system is shown in Fig. 2. The experimental tube was baked for 2 days at 300°C, with traps cooled in dry ice, and subsequently the trap separating the tube from the rest of the system was cooled in liquid nitrogen. After outgassing schedules, and the firing of molybdenum getters, pressures of below 10^{-9} mm Hg were obtained.

5. ION-BOMBARDMENT CHAMBER

During ion bombardment the argon was ionized by a current of electrons flowing from a hot filament through an accelerating grid. A shield at grid potential prevented material, sputtered or evaporated from the filament and its supports, from reaching the crystal face. Despite this precaution, a further undesirable effect was found. Material that was sputtered from a crystal face, deposited on the glass walls of the tube, and there became charged negatively by electrons from the filament. This material was then sputtered back onto the crystal face. For a compound semiconductor, this means that the separate elements can redeposit on the face. It was not always possible to reduce the grid voltage to a value lower than the sputtering threshold for indium and antimony; hence, a metal shield was placed around the crystal to collect stray electrons during ion bombardment. A typical configuration is shown in Fig. 3.

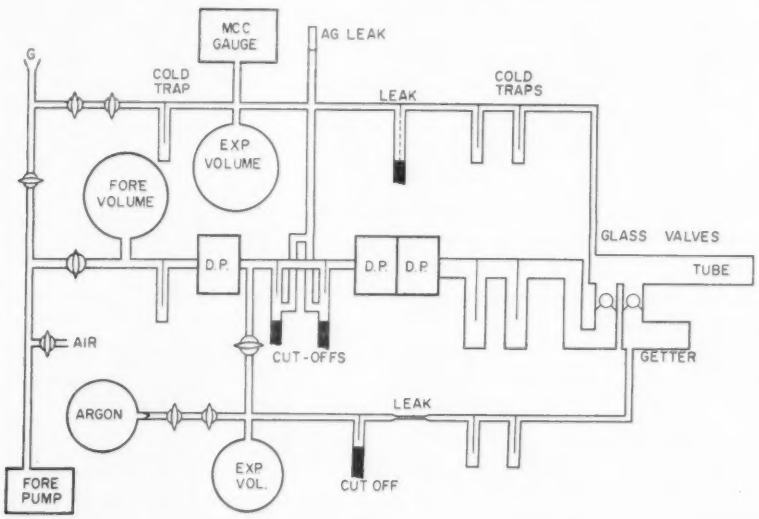


FIG. 2. Schematic diagram of vacuum and gas-handling system.

6. CLEANING PROCEDURES USED

Heating alone of an etched surface for many hours at 350–400°C in high vacuum resulted in very weak diffraction patterns corresponding to an unidentified structure. It was found necessary to bombard the surface with argon ions before heat treatment, in order to obtain diffraction patterns corresponding to a (111) face.

A wide variety of ion-bombardment energies and current densities, combined with various annealing schedules, were tested. The sputtering rate of InSb was found to increase with ion energy, being quite rapid above 200 eV, where it was many times faster than for silicon or germanium. The visual appearance of the bombarded region was detectably altered after sputtering at 400 eV 10 μA/cm^2 for 20 min. The change was due to the formation of small regions having the appearance of black dots under a magnification of 400. The percentage of surface covered by the regions was normally small, and did not appear to contribute to the diffraction patterns, since these were little affected by further sputtering (followed by annealing).

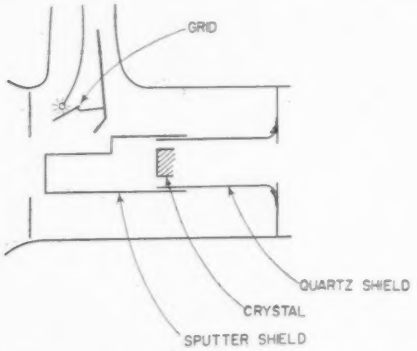


FIG 3. Sketch showing crystal in position for ion bombardment. Note shield to prevent back-sputtering effect.

After a surface was ion bombarded, annealing below 300°C was ineffective in producing diffraction patterns, and below about 330°C it was very slow. The efficiency of annealing appeared greater at higher temperatures, but above about 400°C small hillocks formed on the surface.[6] Hence, standard annealing treatments involved heating the surface for one to several hours at about 350°C, following an ion-bombardment schedule of the order 5 μA/cm^2 at 100–400 eV for 15 min.

7. DIFFRACTION PATTERNS

Reproducible sharp diffraction patterns were obtained from a cleaned (111) surface after several cycles of argon ion bombardment, each followed by annealing at about 340°C for periods of the order of several hours. The intensity of the patterns increased most during the first hour of annealing, after which only a small increase took place during many hours of further heating. The diffraction patterns obeyed the diffraction grating conditions of a surface having an atomic structure expected from the termination of the bulk structure of InSb on a (111) plane. However, half-integral order beams were found, in general, with comparable intensities to the integral order beams, indicating that the surface atoms had double the bulk spacing. A graph of beams positions is shown in Fig. 4.

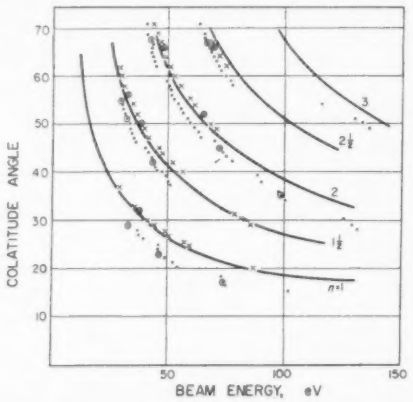


FIG. 4. Positions of diffraction beams in the (110) azimuth for a cleaned (11$\bar{1}$) surface of InSb. Dots (·) represent peaks in plots of collector current vs. beam energy for one (110) azimuth, crosses (×) refer to results from a (110) azimuth at 60° to the above, obtained by rotating the crystal by 60°. The angle between the crystal face and the incident electron beam is slightly different for the two orientations of the crystal as shown by the different deviations of the experimental points from the solid curves. These curves are theoretical plots of the surface grating formula $n = d \sin \theta$, using the bulk lattice constant of 6·48 Å for InSb. Positions at which the beams maximize are represented by a circle around the requisite dot or cross. The order of diffraction, n, is marked on the curves.

The sticking coefficient had a maximum value of about 10^{-5} for oxygen, and was several orders of magnitude lower for nitrogen.

Two types of (111) surface were obtained. In one case an intense diffraction pattern obeying the

(111) surface-grating relationship, and containing both integral and half-integral order beams, was obtained, following a complex treatment, involving overheating and back sputtering of indium and antimony onto the crystal face. Later tests with an omegatron also raised a doubt about the purity of the argon used during this treatment. The sticking coefficient of oxygen was orders of magnitude lower than 10^{-5}. In subsequent experiments on crystals which were not subjected to overheating and back sputtering, and where the argon purity was checked with an omegatron, a reproducible diffraction pattern of the same nature as before, but weaker by an order of magnitude, was obtained. The sticking coefficient of oxygen however was 10^{-5}, similar to that for the $(11\bar{1})$ face. Since the latter pattern was reproducible whereas the former, although more intense was not, and since also the reproducible pattern was obtained from a surface which was not subjected to overheating and back sputtering, using possibly impure argon, it is considered that the latter results are the reliable ones.

VOL. 14 1960

8. OXYGEN ADSORPTION AND CLEAN SURFACE REGENERATION

The effect on the surface of exposure to oxygen was ascertained by noting changes caused in the diffraction pattern. The source of oxygen was a heated, sealed silver tube connected to the system by a metal–glass seal. Oxygen passing through the silver tube into the system was as pure, or purer, than the best obtainable commercially.[7] With all filaments cold, the oxygen was let into the tube continuously through an adjustable porcelain leak, while the system was open to the pumps but with the getter chamber isolated by a ball-and-socket valve. The dynamic equilibrium pressure of oxygen in the tube was estimated from previous experiments using pure nitrogen flowing through the leak. With nitrogen it was possible to use the ionization gauge as a pressure monitor, but this is not possible with oxygen since it forms CO by interaction with carbon on the hot gauge filament.[8]

The adsorption of oxygen resulted in a gradual weakening of all beams in both (110) and (112) azimuths on (111) and $(11\bar{1})$ surfaces. For example, after a cumulative exposure of 5×10^{-3} mm min, the diffraction patterns were reduced to about 25 per cent of their original intensity as seen in Fig. 5. The maximum sticking coefficient* was 10^{-5}. The clean surface could be regenerated by an ion bombardment and anneal cycle, or by annealing only, at about 330°C for about 1 hr or more. If a partially

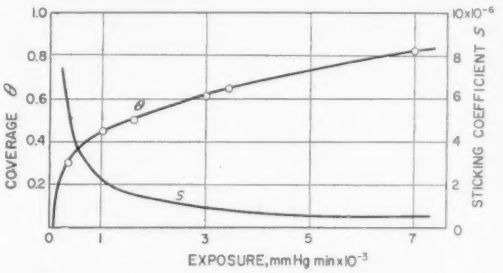


FIG. 5. Fractional coverage θ and sticking probability S as functions of exposure to oxygen at 300°K.

oxygen-contaminated surface was heated at 200°C, the diffraction pattern deteriorated further. However, heating above 320°C caused the diffraction pattern to become more intense.

9. DISCUSSION

(a) *Composition of surface*

When a material containing two kinds of atoms is subjected to ion bombardment, there is a possibility that one kind of atom may be preferentially sputtered, leaving a non-stoichiometric surface. For InSb, under the bombardment conditions used, it is concluded that any tendency for preferential sputtering was sufficiently small that the effect could be removed by a subsequent anneal at about 340°C for approximately one hour. This follows since the surface atoms after this treatment had the InSb structure and symmetry. It is not considered probable that excess In or Sb would take up this structure since the elements are face-centered tetragonal and rhombohedral, respectively, with a quite different symmetry from that of InSb. It may be noted also that In and Sb have similar atomic weights so that very dissimilar sputtering rates are not expected.

At the annealing temperatures used here appreciable surface decomposition does not appear to take place.

* The method of computing the sticking coefficient is given in the author's reply to GREEN and MAXWELL, at the end of the paper.

Table 1. *Low-energy electron-diffraction beams of other than integral order found on cleaned surfaces of* Ge, Si *and* InSb

	Crystallographic surface		
Material	(111)	(110)	(100)
Ge	½ beams, all azimuths,* plus weaker fractional-order beams.	½ beams (110) azimuths,* weaker fractional-order beams in other azimuths.	½ beams in azimuths having even sums of indices.*
Si	fractional-order beams		as above, also fractional-order beams.
InSb	½ beams all azimuths, plus very weak fractional beams		

"½ beams" means beams of order 1/2, 3/2, 5/2, etc.
"Fractional order beams" refers to orders other than multiples of 1/2.
*"½ beams" weakened by O_2 adsorption more than integral order beams.

(b) *Double spacing*

A significant feature of the diffraction results is the presence of half-integral order beams in all azimuths of all faces examined. This feature had also been found for (111) and (110) faces of Ge,[1] and to some extent on (100) faces of Ge and Si. A summary of results for these three semiconductors is given in Table 1.

It appears that there is some degree of generality in the observation of a double-spaced structure on the diamond-structure semiconductors. The phenomenon apparently arises because the atoms in the surface layer have incomplete "bonds". A lower energy equilibrium situation may thus be one where there is a lower density of atoms in a surface plane than in the corresponding bulk plane. A possible atomic configuration for a (100) Ge surface has been suggested by SCHLIER and FARNSWORTH[9]. Further evidence for a double-spaced surface for diamond-structure materials has been deduced from microcleavage characteristics by WOLFF[10].

(c) *Difference between* (111) *and* (11$\bar{1}$) *faces of* InSb

These surfaces are produced by a cut parallel to a (111) plane. When etched in the chemical mixtures described in Section 2, one face, referred to as (11$\bar{1}$), develops etch pits, while the other face, (111), does not. The general surface appearance of the two faces are different after etching.[4] Although this phenomenon, which is found also with other semiconductors in Groups III–V, is not understood in detail it is certainly related to the asymmetry of the zinc-blende type structure in the (111) direction.

From consideration of the bulk structure it is expected that the atomic surface planes themselves are different for (111) and (11$\bar{1}$) faces. On one of these, any antimony atom is triply bonded to the underlying structure, and any indium atom is singly bonded. On the opposite surface the bond numbers are reversed, as seen from Fig. 6. It appears that a break through the single bond, as indicated by the line (1), is energetically more favorable than that shown by the line (2). In this case the atoms on the surface would all be of one

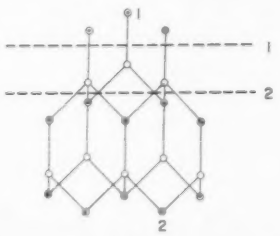


FIG. 6. Appearance of bulk InSb structure, looking along a (112) direction. Alternate atoms are In and Sb. If surfaces are produced by the cuts 1 or 2 along (111) planes, then the same atoms are bonded to the bulk structure differently for the two surfaces.

kind, indium atoms on one surface and antimony atoms on the other. In fact, however, the presence of half-integral order beams in the diffraction pattern indicates that some rearrangement of surface atoms from the bulk positions has taken place, caused at least in part by the presence of unsatisfied surface bonds. The simple concept of the bulk solid ending on either the surface (1) or (2) in Fig. 6, may thus require modification.

The particular modifications to be considered should take into account the following features. (a) As mentioned above, the observation that the surface monolayer has the InSb structure and symmetry, tends to suggest that the surface does not consist of sufficient excess In or Sb to form the observable structure of either of these elements. (b) The observation of a similar (111) type structure for the (111) and (11$\bar{1}$) surfaces indicates a similar atomic arrangement for these surfaces. (c) The values at which the beams maximized were different for the two sides. Since these values are determined in part by the distance between the first and second surface layers, the experimental evidence indicates, though not conclusively, that the spacing of the surface atomic layer from the second layer may be different for the two sides.

The sticking coefficient for oxygen was found to be not appreciably different for the (111) and (11$\bar{1}$) sides having the reproducible diffraction patterns. If a face containing predominantly indium atoms had different oxygen adsorption rates from one containing mainly antimony atoms then this evidence might appear to suggest that both types of atoms were about equally plentiful on (111) and (11$\bar{1}$) surfaces. However, this conclusion need not hold due to the comparatively large distances between atoms in a double-spaced structure. Oxygen atoms might adsorb onto the atoms in a single-spaced second layer exposed through the spaces in the double-spaced surface layer, and thus distort the surface layer itself, leading to a weakening in the observed diffraction pattern. If the first and second layers are composed of indium and antimony atoms, respectively, or vice versa, then both types of atoms would thus take part in the oxygen adsorption process, tending to make the observed adsorption rates for (111) and (11$\bar{1}$) surfaces similar.

The three facts of interest for the surface layers are thus:

(i) the atoms show the InSb structure and symmetry.
(ii) there is a marked tendency for the atoms to show a double-spaced structure,
(iii) the oxygen adsorption rates for (111) and (11$\bar{1}$) surfaces are apparently closely similar.

(d) *Comparison of temperatures for oxygen removal and bombardment annealing*

After argon ion bombardment, it has been found that InSb, like germanium and silicon and other crystals, must be annealed at a certain temperature T_1 to remove the lattice damage caused by the bombardment. For these three semiconductors it has also been found that the diffraction pattern from a clean surface can be restored after oxygen adsorption by simply heating the surface to a certain temperature T_2. These temperatures are given in the following table.

Table 2

Material	Anneal temp. T_1	Regen. temp. T_2	m.p.
Ge*	600°C	500°C	958°C
Si*	1000°C	900°C	1420°C
InSb	340°C	330°C	523°C

* From SCHLIER and FARNSWORTH[1].

It is seen that there is a marked similarity between T_1 and T_2 for these semiconductors. This can be considered in relation to the mechanism by which oxygen is removed from a heated surface. For germanium, the heat of adsorption of oxygen is of order 100 kcal/mole so that desorption at only 500°C would appear to be negligible. For InSb, the heat of adsorption of oxygen[11] is of order 2 kcal/mole so that some desorption at 330°C is perhaps possible, though the temperature is rather low. Another mechanism for oxygen removal from germanium has however been suggested by FARNSWORTH[9] whereby oxygen diffuses into the bulk. This process is presumably aided when the atoms of the material attain sufficient energy to move back into equilibrium positions after the ion bombardment. That is, the mobility required by the surface atoms to recover from the nonequilibrium positions caused by ion bombardment is also that needed to permit a ready

passage of oxygen atoms between them. This process may not correspond with the bulk-diffusion rate of oxygen.

The above mechanism for oxygen removal has been suggested by the correspondence between the temperatures T_1 and T_2 for the three materials discussed. It is of course possible that there is a degree of coincidence in this, and that quite different oxygen removal mechanisms are important.

Acknowledgement—The author is grateful to Professor H. E. FARNSWORTH for critical and valuable discussions of this work.

REFERENCES

1. SCHLIER R. E. and FARNSWORTH H. E., *J. Chem. Phys.* **30,** 917 (1959).
2. FARNSWORTH H. E., SCHLIER R. E., GEORGE M. and BURGER R. M., *J. Appl. Phys.* **26,** 252 (1955).
3. HANEMAN D., *J. Phys. Chem. Solids,* **11,** 205 (1959).
4. LAVINE M. C., ROSENBERG A. J. and GATOS H. C., *J. Appl. Phys.* **29,** 1131 (1958).
5. FARNSWORTH H. E., SCHLIER R. E., GEORGE M. and BURGER R. M., *J. Appl. Phys.* **29,** 1150 (1958).
6. HANEMAN D., *J. Appl. Phys.*, in press (1959).
7. WHETTEN N. R. and YOUNG J. R., *Rep. 19th Annu. Conf. Phys. Electronics,* p. 147. Massachusetts Institute of Technology, Cambridge (1959).
8. SCHLIER R. E., *J. Appl. Phys.* **29,** 1162 (1958).
9. SCHLIER R. E. and FARNSWORTH H. E., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 3, University of Pennsylvania Press, Philadelphia (1957).
10. WOLFF G. A., *Acta Cryst. Camb.* **12,** 313 (1959).
11. ROSENBERG A. J., LAVINE M. C. and MARTEL C. S., *Quart. Prog. Rep., Lincoln Lab.* p. 13 (Aug. 1958).

VOL.
14
1960

DISCUSSION

J. J. LANDER (Bell Laboratories): In that part of the paper indicated as Section 9(a), Composition of the Surface, the author states that "it is not considered probable that excess In or Sb would take up this structure" (that of InSb) "since they are face-centered tetragonal and rhombohedral, respectively..." and he concludes that the surface is stoichiometric. However, evidence from studies of epitaxial growth with many combinations of materials show that the structure of the first few atom layers of foreign atoms deposited on surfaces may be radically influenced by the base material. Thus a monolayer of chemisorbed In or Sb would be most apt to take up sites provided by InSb.

A second comment concerns the possibility of precipitation of a solute with relatively low solubility. In experiments on surfaces which follow an anneal of some intermediate temperature one must keep in mind that if the bulk material at a higher temperature was effectively saturated with a solute, then this may precipitate on the surface during the anneal. Recent studies of precipitation on dislocations indicate the prevalence of this comparable phenomenon. Solutes with very low solubilities may be active. Examples with ZnO as the solute are discussed in the twenty-second paper of this meeting.

REPLY: In our experiments the processes which take place are not deposition of indium or antimony on the surface but *removal* of these materials. If one element is removed preferentially during sputtering and or heating, then the surface region might contain an excess of the other element. Such In- or Sb-rich regions would tend to segregate after annealing into InSb plus the excess element.* However the diffraction patterns obtained after the prescribed treatments, were characteristic of InSb indicating that any other structures at the surface were either disordered or present in negligible amounts.

The possibility that impurities may precipitate at the surface during heating cannot be excluded, even though the specimens used were as pure as obtainable. However, after a cleaned surface pattern was obtained, further heating for up to 35 hr at around 350°C (the bombardment-anneal temperature region) did not appreciably affect the diffraction beams or oxygen adsorption characteristics. This would indicate that at the temperatures used here any possible tendency for impurities to precipitate at the surface caused less than about 5 per cent change in surface composition even after 35 hr of heating.

M. GREEN and K. H. MAXWELL (Zenith Radio Corporation): In their work on germanium surfaces SCHLIER and FARNSWORTH found that certain of the diffraction beams were very nearly extinguished by the adsorption of a monolayer of oxygen atoms. The amplitude of such a beam was assumed to be proportional to the uncovered surface area. If I_0 is the intensity of the line for a bare surface, then the intensity I at coverage θ was assumed to be given by

$$I = I_0(1-\theta)^2.$$

Since the adsorption of oxygen on the InSb surface resulted in a gradual weakening of *all* the beams, with no pattern characteristic of the oxygen covered surface, how were the values of θ shown in Fig. 5 derived? In particular, at what point in the weakening of the diffraction pattern by oxygen adsorption was it assumed that θ was unity?

By what method was the sticking coefficient, S, calculated? If it was calculated by the method used by SCHLIER and FARNSWORTH then, as one of us has shown,† there is still doubt as to the correct analysis of the experimental results.

Since the derivation of a value for the sticking coefficient, S, depends on a knowledge of the absolute pressure, could an estimate of the error in pressure measurement be given?

In view of the above points, what significance can be attached to the reported values of θ and S?

REPLY: In obtaining the values of θ and S the following assumptions and steps were used:

(1) The diffracted beam intensity is taken proportional to the uncovered surface area. The fractional coverage θ then equals unity when the beams are extinguished.

(2) From the values of θ obtained as above, the sticking coefficient S may be computed by the method used by SCHLIER and FARNSWORTH as given in a more recent publication‡ than the one to which GREEN refers. Since the number of molecules that strike the surface during a given exposure is known, S may be calculated directly from the measured variation of θ with exposure:

$$S = N[d\theta/d(\int p\,dt)]$$

where $N = N_s/(4{\cdot}3\times10^{22})$ being the ratio of the number of atoms/cm^2 adsorbed at full coverage, to the number of oxygen atoms striking 1 cm^3 at 300°K during an exposure of 1 mm Hg-min.

(3) The value of N_s was taken as the number of atomic sites on an undistorted (111) plane of InSb.

Concerning pressure measurements, the equilibrium oxygen pressures in the tube were obtained from the settings of a calibrated porcelain leak, as discussed in the text. Tests of the calibration with gases N_2 and CO indicated an upper limit to the error in oxygen pressure measurements of about 25 per cent.

---

* BUSCH G., *Nuovo cim. Suppl.* **7,** 696 (1958).

† GREEN M., *Progress in Semiconductors*, Vol. 4, p. 35. Heywood, London, (1960).

‡ SCHLIER R. E. and FARNSWORTH H. E., *J. Chem. Phys.* **30,** 917 (1959).

In assessing the significance of the values of θ and S, the main possibility of error in computing S would appear to lie in (3). If the oxygen tends to deposit on the surface in small clumps the numerical value of N_s will be affected, by as much as a factor of 2 or 3 in extreme cases.

The other assumptions involved are of course not above controversy due to present insufficient knowledge of the precise processes involved in adsorption. Taking everything into account however, the quoted values are nevertheless expected to be within a small integral factor of the true values.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 169–174. Printed in Great Britain.

ETCHING AND INHIBITION OF THE {111} SURFACES OF THE III–V INTERMETALLIC COMPOUNDS: InSb*

H. C. GATOS and M. C. LAVINE

Lincoln Laboratory, Massachusetts Institute of Technology, Lexington 73, Massachusetts

Abstract—In oxidizing etching media, the {111} surfaces of InSb are much less reactive than the {$\bar{1}\bar{1}\bar{1}$} surfaces as determined by direct dissolution rates and by etching of spheres. Related to this difference in reactivity is the fact that dislocation etch pits appear on the {111}, but not on the {$\bar{1}\bar{1}\bar{1}$} surfaces. Inhibitors such as stearic acid and primary amines decrease appreciably the dissolution rate of only the {$\bar{1}\bar{1}\bar{1}$} surfaces. Furthermore, etching inhibitors lead to the formation of new dislocation etch pits on the {111} and also on the {$\bar{1}\bar{1}\bar{1}$} surfaces. Similar effects were observed in GaAs. These phenomena are discussed in terms of the atomic configuration of the surfaces involved and the relative reactivity of the group III and group V atoms.

VOL. 14 1960

INTRODUCTION

WITH the introduction of intermetallic compounds to semiconductor science and technology, surface studies of these materials have become of interest. The surfaces of the III–V compounds (zinc-blende structure) present unusual characteristics as a result of the crystallographic polarity along the ⟨111⟩ direction. It is well known, for example, that dislocation etch pits can ordinarily be revealed on the {111} surfaces of these compounds but not on the {$\bar{1}\bar{1}\bar{1}$} surfaces. WAREKOIS and METZGER[1], employing X-ray diffraction techniques, have recently identified the {111} surfaces terminating with group III atoms and those terminating with group V atoms.† It is now generally accepted that the {111} surfaces developing dislocation etch pits are the A surfaces, whereas those developing no dislocation etch pits are the B surfaces.

The authors[2] have recently discussed these and other differences between the two types of {111} surfaces. They have also reported on the influence of the polar ⟨111⟩ direction on the etching behavior of the {110} and {100} surfaces of InSb.[3] In these reports an attempt was made to interpret the observed etching behavior on the basis of bonding and electronic configuration.

In the present paper the differences between the A and B surfaces of InSb are pursued further, and new effects of etching inhibitors leading to the development of dislocation etch pits on both A and B surfaces are presented. An interpretation of these results is advanced.

EXPERIMENTAL

The specimens employed in this study were cut from high-purity single crystals of InSb grown in the ⟨111⟩ direction by the Czochralski method (rate of rotation: 10 r.p.m.; rate of pulling: 1·9 cm/hr). The parallel sides of wafer specimens were ground to within 1° from the {111} orientation as determined by X-ray diffraction. Regular tetrahedrons (with an edge of approximately 1 cm) having either only A or only B faces were also prepared by grinding. In InSb[4] as in the case of other III–V compounds the A and B surfaces have been identified by X-ray diffraction techniques.

The InSb spheres (approximately 1 cm in diameter) were prepared by placing specimens of roughly cubical shapes on a tube mounted on a horizontal wheel. The specimens were rotated

* The work reported in this paper was performed by Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology with the joint support of the U.S. Army, Navy and Air Force.

† In the III–V compounds it is customary to designate the group III atoms as A atoms and the {111} surfaces terminating with triply bonded A atoms as A surfaces. Similarly the group V atoms are designated as B atoms and the {111} surfaces terminating with triply bonded B atoms as B surfaces. This designation will be used in the present paper.

manually in a direction opposite to the rotation of the wheel.*

The two principal etchants employed were: (a) 2 conc. HNO_3 + 1 conc. HF + 1 glacial CH_3COOH (in parts by volume); and (b) 0·2N Fe^{+++} in 6N HCl. For convenience, etchant (a) will be referred to as modified CP–4 and (b) as Fe^{+++} etchant. In the case of InSb, the modified CP–4 generally leads to chemical polishing whereas the Fe^{+++} etchant acts preferentially leading to well defined etch figures.

Etching experiments

Dissolution rates. At room temperature, the dissolution rate of InSb in strongly oxidizing media, including the modified CP–4 and the Fe^{+++}

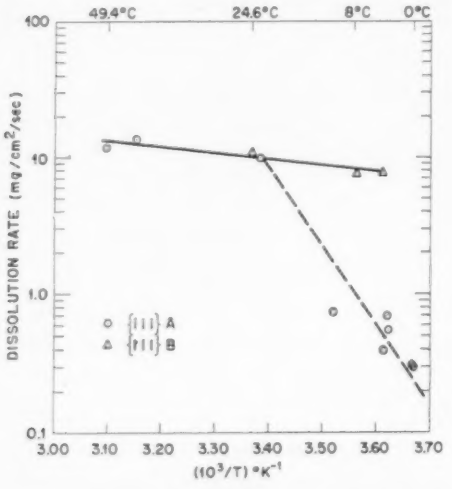


FIG. 1. Dissolution rate of InSb as a function of temperature in modified CP–4.

etchant, is controlled by the diffusion rate of the oxidizing species to the InSb surface. The apparent activation energy for dissolution is approximately 5 kcal/mole, typical of diffusion controlled processes. Consequently, no differences in dissolution rates were observed between the A and B tetrahedrons. At lower temperatures however, differences in dissolution rate exceeding one order of magnitude were observed, as shown in Fig. 1. Here, the dissolution rate of the A surfaces becomes a chemical activation process with an apparent activation energy of approximately 25 kcal/mole. The dissolution of the B surfaces remains under diffusion control down to, at least, 0°C.

Etching characteristics. Although, at room temperature, no difference in dissolution rates was observed between the A and the B surfaces, the microstructure of the A and the B surfaces etched in modified CP–4, was found different. For example, the familiar dislocation etch pits were observed on the A but not on the B tetrahedrons.

In the Fe^{+++} etchant, the etch figures of the A and the B surfaces were strikingly different. The etch figures on the B surfaces formed arrays of well defined or overlapping equilateral triangles oriented along and bounded by ⟨110⟩ directions. The etch figures on the A surfaces were irregular or regular hexagons usually much larger and less well defined than the corresponding triangular figures of the B surfaces. The hexagons were also bounded by ⟨110⟩ directions. A typical example of the etch figures on the A and B surfaces of an {111} wafer is shown in Fig. 2.

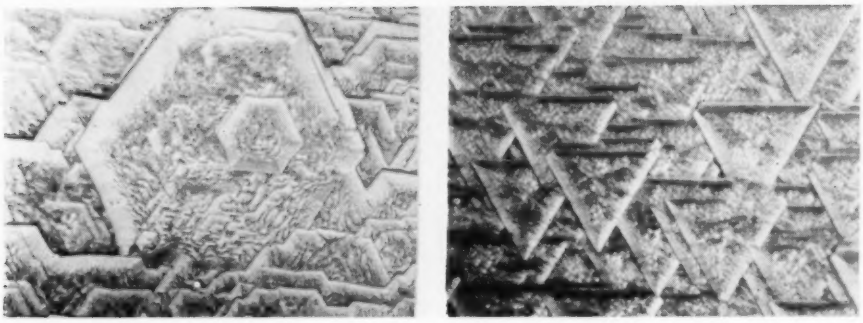
The above differences in etch figures can serve for the unmistakable identification of the A and B surfaces even when, due to very low dislocation densities, dislocation etch pits are not revealed on the A surfaces of InSb. In this connection it is of interest to point out that triangular figures (not as well defined as in Fig. 2) persist on the B surfaces even when the A surfaces appear chemically polished.

Etching of spheres. Single crystal spheres can provide a simple means for obtaining general information on the orientation dependence of etching. This is true even if the overall etching rate is controlled by transport phenomena in the etchant.

InSb spheres were etched in several media and at various temperatures. In no case did the A faces develop since they are the slowest reacting ones. Consequently, octahedron shapes resulting from the development of all of the {111} faces, as in the case of germanium,[5] were never observed. The orientation of the spheres was determined by X-ray diffraction and preferential etching techniques. Figure 3(a) shows a sphere etched in modified CP–4 at room temperature. The pitted triangular area corresponds to a B face. Four such

* This method was suggested to us by G. A. WOLFF of the U.S. Army Signal Research and Development Laboratory.

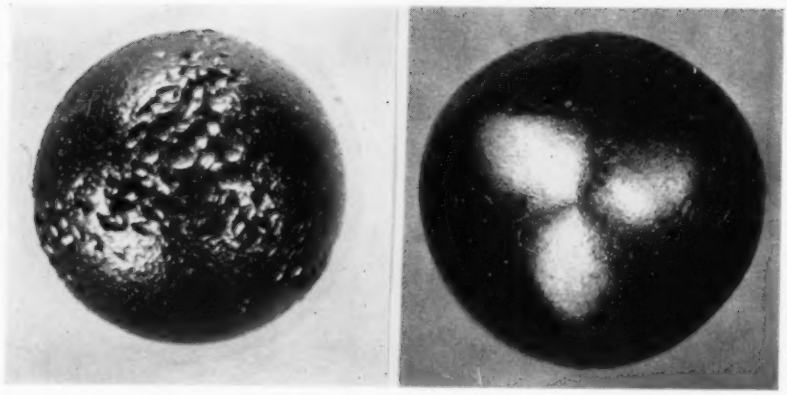
VOL.
14
1960



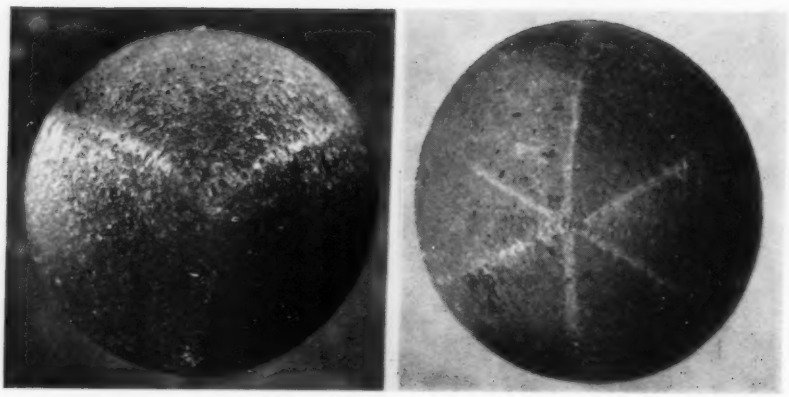
(a) (b)

FIG. 2. The A (a) and the B (b) surface of an InSb wafer etched in Fe^{+++} etchant for 30 min at 87 °C × 315.

[*facing p.* 170



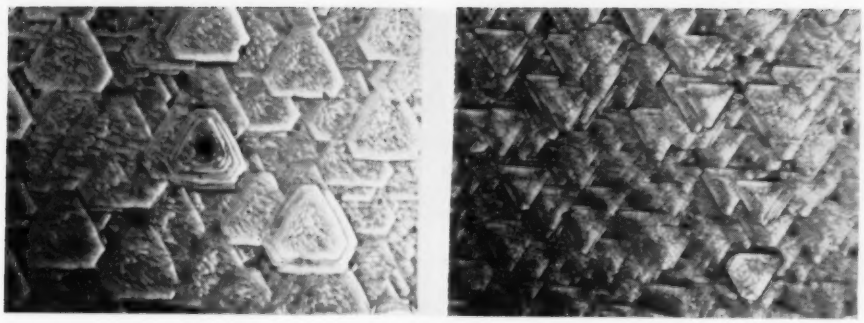
(a) (b)



(c) (d)

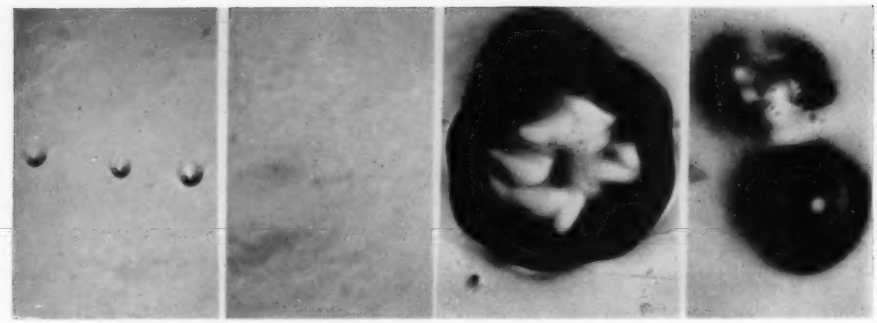
FIG. 3. (a) InSb etched at room temperature in 1 part modified CP–4+1 part H_2O (by volume); the pitted triangular area corresponds to a B face. (b) Sphere etched in 1 part 30 per cent H_2O_2+1 part conc. HF+4 parts H_2O (by volume), showing one of the four three-{100}-face sets. (c) Sphere etched in Fe^{+++} etchant at 80°C, showing a three -{100}-face set merging into a B face. (d) Sphere (c) showing one of four six-face sets with three {110}-and three {100}-faces. Spheres were approximately 1 cm in diameter.

VOL.
14
1960



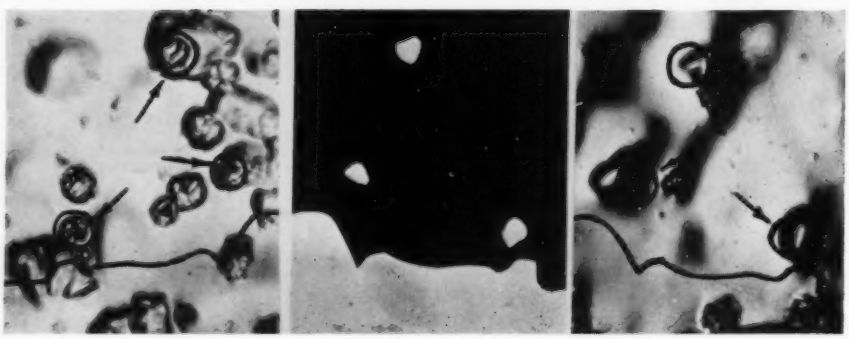
(a) (b)

FIG. 4. The A (a) and the B (b) surface of an InSb wafer etched for 30 min at 87°C in the Fe^{+++} etchant containing 0·5 per cent butylamine × 315.



(a) (b) (c) (d)

I



(a) (b) (c)

II

FIG. 5. I (a) and I (b) are the A and B surfaces, respectively, of an InSb wafer etched for 1 min in modified CP–4 at 0°C; 315 ×. I (c) and I (d) are the A and B surfaces, respectively, etched for 1 min in modified CP–4+0·5 per cent butylamine at 0°C. The decrease in size of the conical pits can be noted in going from I (a) to I (c); 315 ×. II (a) and II (c) are the A and B surfaces, respectively, of a wafer etched in modified CP–4+0·5 per cent amylamine; 100 ×. Subsequent etching of same wafer resulted in formation of holes as shown in II (b) (thickness decreased by approximately 0·3 mm); holes and outline of II (b) are indicated in II (a) and II (c) for comparison.

areas developed around the sphere and were connected by pitted strips corresponding to the {100} planes. Independent tests showed that in the modified CP–4 etchant the {100} surfaces are approximately as reactive as the B faces.

In 1 part 30 per cent H_2O_2 + 1 part conc. HF + 4 parts H_2O (by volume), in addition to the four B faces, four sets of three {100} faces were observed. One of these sets is shown in Fig. 3(b). The faces of each set converge to an A {111} pole. In the Fe^{+++} etchant each of the four B faces were formed at the intersection of three {100} faces (Fig. 3(c)); in addition, four sets of six faces were developed (one such set is shown in Fig. 3(d)); three of the six faces are {100} and three are {110}. The six faces converge to an A {111} pole.

VOL. 14 1960

Inhibition experiments

Some of the etching experiments described above were repeated in the presence of etching inhibitors in the hope that etching inhibition might contribute to the understanding of the etching mechanisms. Tetrahedral specimens, as above, were employed for determining the dissolution rates.

The following inhibitors were employed: stearic acid, amylamine, butylamine and propylamine. About 0·5 per cent of a given amine was added to the etching solution. Since the solubility of stearic acid is very small, a slight excess of this inhibitor was added to the etchants.

In the modified CP–4 at O°C containing any of these inhibitors, the dissolution rate of the A surfaces was found to be $0{\cdot}31 \pm 0{\cdot}04$ mg/cm²/sec; a rate of approximately $0{\cdot}52 \pm 0{\cdot}08$ mg/cm²/sec was obtained in the absence of inhibitors. In the case of the B surfaces, the dissolution rate decreased very markedly in the presence of an inhibitor; i.e. from approximately 7·8 mg/cm²/sec to $0{\cdot}36 \pm 0{\cdot}04$ mg/cm²/sec.

The fact that the dissolution rates of the A and B surfaces became virtually equal in the presence of inhibitors was also reflected in the etching of spheres; their shape was not altered detectably even after prolonged etching. Thus, although no independent experiments were performed, it would appear that the dissolution rates of the {110} and {100} surfaces are inhibited to approximately the same values found for the A and B surfaces.

It is of interest to note that the differences in microstructure between the A and B surfaces (Fig. 2) become far less pronounced in the presence of inhibitors as seen in Fig. 4.

Dislocation etch pits on the A *and* B *surfaces.* The most striking effect of the etching inhibitors is the appearance of dislocation etch pits on the B surfaces and the appearance of new pits on the A surfaces. A preliminary account of this observation has already been reported.[6] The etch pits ordinarily appearing on the A surfaces appear also in the presence of inhibitors, but their size is appreciably smaller than that observed under the same conditions in the absence of inhibitors. Typical examples of an A and a B surface etched in modified CP–4 with and without an inhibitor are shown in Fig. 5–I (all of the inhibitors pointed out above lead to the same results).

The following experiment was performed in connection with the etch pits developing in the presence of inhibitors: An {111} wafer approximately 3 mm thick was etched in modified CP–4 containing no inhibitor. The A and B surfaces were replicated on Faxfilm and then the wafer was re-etched (without grinding) in modified CP–4 containing an inhibitor. The A and B surfaces were again replicated. Etching in modified CP–4 containing an inhibitor and replicating were repeated nine times until microscopic holes appeared through the wafers. The thickness of the wafer was monitored with each etching. Large areas of the surface replicas were mapped on composite microphotographs using suitable reference points.

In the successive etchings, it was possible to follow through the wafer many of the new pits appearing on the A and B surfaces. In these cases, the pits followed a line intersecting the {111} surface at an angle of 60°, characteristic of dislocation lines in the zinc-blende structure. (An error of ±2° was associated with these determinations.) In many instances it was possible to find one-to-one correspondence between the new pits on the A and those on the B surface. Furthermore, some of the holes finally formed on the wafer, could be traced to individual pits (Fig. 5–II). No direct correspondence was found between the ordinary etch pits on the A surface and new pits on the B surface.

Although only a limited number of the new pits

was traced through the wafer, there is no doubt that they are associated with dislocations.

It should be pointed out that the appearance of new pits in the presence of inhibitors was observed not only in the modified CP–4, but in a number of other etching media. Furthermore, by adding inhibitors in the GaAs etching employed by SCHELL[7] (10 ml conc. HNO_3+30 ml H_2O), etch pits appeared on both A and B surfaces of a GaAs wafer.

DISCUSSION

In a previous communication[2] an atomic model was proposed to explain the overall differences in behavior between the A and the B surfaces of the III–V compounds. The model is based on the assumption that the sp^3 tetrahedral bonding of these compounds extends to the surface atoms. The B atoms on a B surface must then have a dangling (unshared) but filled s–p orbital (Fig. 6(b)). In the case of the A surface atoms, (Fig. 6(a)), there

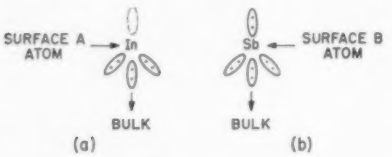


FIG. 6. Atomic model of the A and B {111} surfaces.

are no electrons available for the fourth s–p orbital. A reasonable consequence of this is that the electronic configuration of the A surface atoms is distorted from the tetrahedral symmetry. On this basis it was anticipated and experimentally verified that less perfect crystals of InSb are grown in the A{111} direction than in the B{111} direction.[8]

Regarding the chemical reactivity of the A and B surfaces, it was argued that in reactions where oxidizing (electrophilic) agents are involved B atoms are more reactive than A atoms because they have an unshared pair of electrons available for oxidation.

The dissolution process under consideration is far too complex and the available information rather limited to allow a quantitative development of the dissolution mechanism. Certain aspects of the mechanism, however, can be pointed out at this time.

The growth and dissolution of crystals are generally considered to proceed through the propagation of atomic or molecular steps along the exposed surfaces.[9–11]. In the case of the III–V compounds, such steps are expected to be diatomic layers as shown schematically in Fig. 7. In other words, the

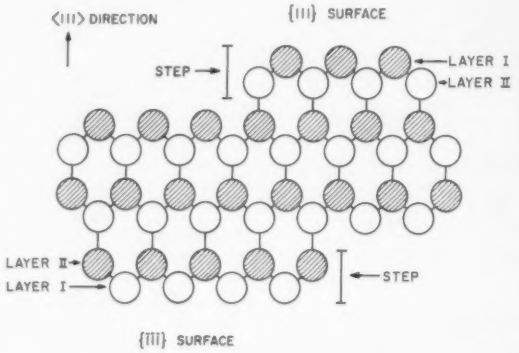


FIG. 7. Two-dimensional representation of the zinc-blende structure with surface steps.

surface configuration with singly bonded atoms is considered unstable; if it were stable, the A and B surfaces would behave alike. Thus, the rate determining (slow) partial reaction* of the dissolution process is not associated with the breaking of the single bonds between the atoms of layers II and the substrate, but primarily with the bonding between layers I and II. This conclusion is consistent with the atomic model of the {111} surfaces pointed out above. It is being assumed, of course, that no drastic structural changes occur at the surfaces, since such changes are not readily visualized at this time.

The rate of motion of the diatomic steps (v_s) across the surface is an important parameter in the dissolution process.[10, 11] Atomic pits (created by the removal of one atom from layer I) or kinks (created for example by the removal of two adjacent atoms from layer I), such as shown in Fig. 8, can be considered as nucleation points for steps (no reference is made to crystal edges, since the crystal surfaces are assumed to be relatively large). If, over a given {111} surface, the rate of

* The term "partial reaction" will be used in lieu of "step" from the chemical kinetics standpoint. The term "step" will be retained for designating a physical step as in Fig. 7.

formation of atomic pits or kinks (v_n) is relatively large ($v_n > v_s$), then it would lead to the unstable surface configuration with singly-bonded atoms. Thus, it is necessary that $v_s \geqq v_n$. No reliable conclusion can be drawn as to the relative magnitudes

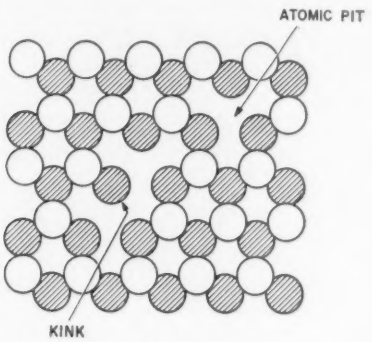


FIG. 8. Representation of an {111} surface with an atomic pit and a kink.

of v_s and v_n. However, in view of the fact that the dissolution rate of the B surface is much greater than that of the A surface, it can be stated that

$$v_n^b > v_n^a$$

where the superscripts a and b refer to A and B surfaces. This relationship probably reflects the marked tendency of the B faces to develop well defined etch figures.

Etching at dislocations

There are two types of edge dislocations present in the III–V compounds; i.e. α dislocations, having a row of A atoms and β dislocations, having a row of B atoms. VENABLES and BROUDY[12] have shown that on the {110} and the {111} surfaces of InSb the ordinarily observed dislocation etch pits are associated with only one type of dislocation. If their work is viewed in the light of WAREKOIS'[4] identification of the InSb crystallographic polarity, the observed dislocation etch pits are associated with α (In) dislocations. Such dislocations in intersecting the A surfaces terminate with a doubly bonded surface atom. This atom is apparently attacked faster than the other surface atoms. Thus, it is reasonable to expect that the rate of nucleation (v_d^α) along the dislocation line (exposing at any given time a doubly bonded atom) exceeds v_s^a, which is a condition necessary for the formation of a dislocation etch pit. On the other hand, β dislocations in intersecting the A surfaces terminate with a trivalent atom. Thus, it is not surprising that the rate of nucleation along the β dislocation (v_d^β) does not exceed v_s^a. The elastic strains associated with edge dislocations do not appear to play a significant role in the etching of the III–V compounds.

Regarding the B surfaces, v_s^b is apparently so large that it is not ordinarily exceeded by v_d^α or v_d^β.

Inhibition

In acid solutions, the primary amines employed as inhibitors form positively charged ammonium ions and are expected to adsorb preferentially (although not exclusively) on the B surface atoms, which have an unshared pair of electrons (see also model Fig. 6). It is believed that this is the primary reason for the marked decrease in dissolution rate of the B surfaces in the presence of inhibitors and the virtual lack of inhibition in the case of the A surfaces. Preferential adsorption on the B atoms can also be argued for the case of stearic acid. Although the substantial decrease in v_n^b is apparent, it is not clear to what extent v_s^b is affected. The fact that the surface dimensions of the α dislocation etch pits on the A surface decrease (but not their depth) suggests a decrease in v_s^a, although the overall dissolution rate of the A surfaces is only slightly decreased.

Regarding the new dislocation etch pits developing on both the A and B surfaces in the presence of inhibitors, their formation is apparently related to a decrease in v_s^a and v_s^b. These new etch pits can be associated with β dislocations on the A surface and with α or β dislocations on the B surfaces. There are indications that some of the new pits may be associated with screw dislocations. Any further discussion on the formation of these new pits must be postponed until more is known about the nature of the dislocations with which they are associated.

Acknowledgements—The authors are indebted to H. H. EHLERS and S. C. ROWE for their assistance with the experiments and to T. P. TURNBULL for his assistance with the sample replicas.

REFERENCES

1. WAREKOIS E. P. and METZGER P. H., *J. Appl. Phys.* **30,** 960, (1959); see also WHITE J. G. and ROTH W. C., *J. Appl. Phys.* **30,** 946 (1959).
2. GATOS H. C. and LAVINE M. C., submitted to *J. Electrochem. Soc.*
3. GATOS H. C. and LAVINE M. C., submitted to *J. Electrochem. Soc.*
4. WAREKOIS E. P., unpublished results.
5. ELLIS R. C., Jr., *J. Appl. Phys.* **25,** 1497 (1954).
6. GATOS H. C. and LAVINE M. C., *J. Appl. Phys.*, in press.
7. SCHELL H. A., *Z. Metallk.* **48,** 158 (1957).
8. GATOS H. C., MOODY P. L. and LAVINE M. C., *J. Appl. Phys.*, in press.
9. BURTON W. K., CABRERA N. and FRANK F. C., *Phil. Trans.* **243,** 299 (1951).
10. CABRERA N., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 327. University of Pennsylvania Press, Philadelphia, 1957.
11. GILMAN J. J., JOHNSON W. G. and SEARS G. W., *J. Appl. Phys.* **29,** 747 (1958).
12. VENABLES J. D. and BROUDY R. M., *J. Appl. Phys.* **29,** 1025 (1958).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 175–180. Printed in Great Britain.

THE OXIDATION OF INTERMETALLIC COMPOUNDS—III
THE ROOM-TEMPERATURE OXIDATION OF $A^{III}B^{V}$ COMPOUNDS*

A. J. ROSENBERG

Lincoln Laboratory, Massachusetts Institute of Technology, Lexington 73, Massachusetts

VOL. 14 1960

Abstract—The oxidation of vacuum-crushed InSb, GaSb, AlSb, InAs, GaAs and InP has been studied at 26°C, and the kinetics have been analyzed in terms of the empirical function, $\eta \equiv -\partial \log V/\partial X$. (The general utility of η and the associated function, $\epsilon = \eta X$, in the kinetics analysis of protective surface reactions is discussed in detail.) Maxima in η are observed at $2{\cdot}5 \times 10^{15}$ and $1{\cdot}35 \times 10^{15}$ oxygen atoms/cm^2 on the antimonides and the arsenides, respectively. The reaction of InP exhibits no distinct maximum in η; otherwise, it resembles that of the arsenides. The results are interpreted in terms of metastable surface complexes whose formation precedes the initiation of true oxide growth. The composition of these intermediates is dictated largely by the nature of the B^{V} atom in the substrate. The associated kinetics depend systematically, however, upon the A^{III} atoms. Thus, the rate of complex formation decreases in the order Al > In > Ga, while the rate of oxide formation decreases in the order In > Ga > Al.

INTRODUCTION

THE unusual physical and chemical properties of the $A^{III}B^{V}$ semiconducting intermetallic compounds are the subject of growing scientific and technological interest. The earlier papers in this series[1, 2] considered the effects of these properties upon the high temperature (212°–494°) oxidation of the representative compound, InSb. The present paper is concerned with the reactions of oxygen with InSb, GaSb, AlSb, InAs, GaAs and InP at room temperature. Evidence is presented for the formation of metastable intermediate complexes whose structure depends primarily upon the B^{V} atoms but whose stability depends upon the A^{III} atoms. The results also provide a decisive kinetic delineation between the formation of the complexes and the subsequent initiation of true oxide structures.

EXPERIMENTAL PROCEDURE

Pure single crystals were crushed and apportioned to sample chambers appended to a sealed all-glass apparatus which had been baked out under high vacuum. The essential features of this procedure have been described previously.[3, 4] It has been shown[5] in another study dealing with the oxidation of germanium that the outgassing of the glass walls after the prolonged vacuum bakeout is insignificant compared to the monolayer capacity of the crushed samples. The sample chambers, containing ca. 0·5g of material with 1000–3000 cm^2 of fresh surfaces, were sealed off from the crushing apparatus and connected through break-seals to a gas adsorption apparatus of a type described elsewhere.[6] The temperature of the sample chambers was maintained at $26{\cdot}2 \pm 0{\cdot}2$°C. Oxygen uptake was measured by following pressure changes at constant volume from the time of breaking the seal to as long as 5 days. The use of a thermistor manometer[6] permitted continuous pressure measurements with a precision of one part in 10,000. With this precision, the absolute rate at which the pressure decreased during a run could be restricted to small values by adjusting the free volume in the system. The occasional admission of additional oxygen into the system sufficed to maintain the pressure between 400–450 μ.

At the conclusion of each experiment the surface area was determined through measurements of krypton adsorption at −195°C.[6] A precise description of each krypton isotherm was obtained through the B.E.T. equation[7] with $C > 100$. The

* The work reported in this paper was performed by Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology with the joint support of the U.S. Army, Navy and Air Force.

surface areas were calculated using $\sigma_{Kr} = 19{\cdot}4$ Å;[6] the absolute values for the surface areas may be as much as 20 per cent in error, but the uncertainty in the relative values is less than 1 per cent.

RESULTS

Since the rate of oxygen sorption decreases by as much as 5 orders of magnitude during a run, the most convenient expression of the data is a plot of N vs. log t where N is the uptake of oxygen atoms normalized to unit surface area and t is the reaction time. The results are so summarized in Fig. 1. Each curve is drawn through more than fifty experimental points, the relative uncertainty in each point being less than the thickness of the curve.

No measureable quantity of oxygen could be desorbed by interrupting the reactions and pumping at 26°C. The sorption is, therefore, essentially irreversible.

DISCUSSION

General features of the reaction

In contrast to the behavior of germanium and silicon,[3] which are shown for comparison, in Fig. 1, the rates are not logarithmic; i.e. $N \neq A + B \log t$. Except for InP, each curve is characterized by a region with negative curvature followed by a broad transition into a region of positive curvature. The inflection clearly signals the onset of a new stage of reaction. Aside from this, it is also evident that the behavior of the antimonides is distinct as a class from that of other compounds.

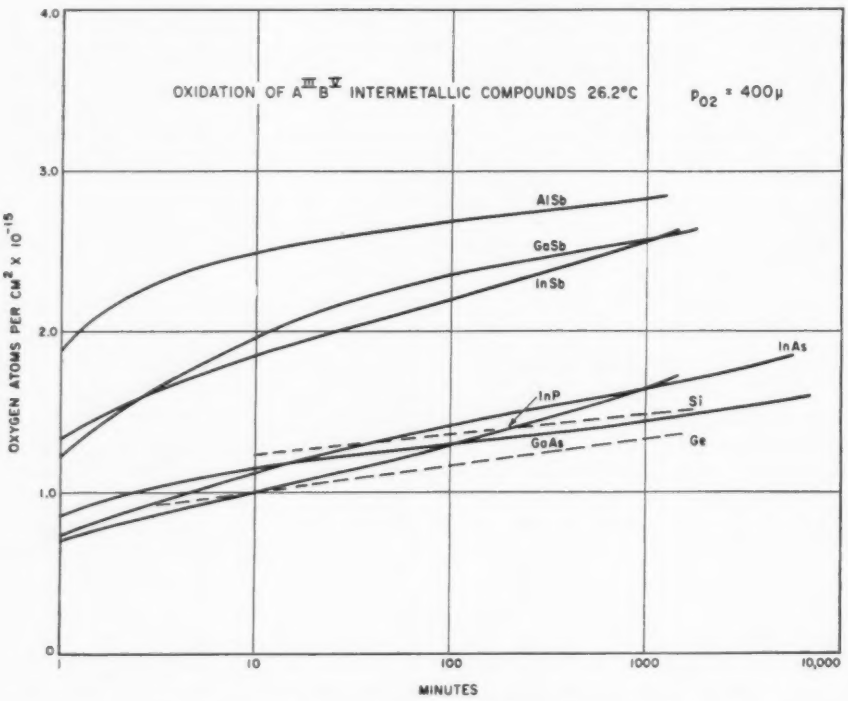


FIG. 1. Kinetics of the oxidation of $A^{III}B^{V}$ compounds. 26·2°C, oxygen pressure = 0·40–0·45 mm.

As will be shown below, the average surface atom density on the surfaces of crushed $A^{III}B^{V}$ compounds is approximately 7×10^{14} cm^{-2}. Unless the curvatures reverse at times much less than 1 min, the surface coverage of each compound will be well under a monolayer of oxygen atoms in reaction times less than 1 sec. This implies that the sticking coefficient of oxygen is very small,

Table 1. Characteristics of the η and ϵ functions

Reaction	$X(t)$	$\eta \equiv -(\partial \log V/\partial X)$	$\epsilon \equiv -(\partial \log V/\partial \log X) = \eta K$
Non-protective:			
auto-catalytic	—	<0	<0
linear	$X = A+Bt$	0	0
Protective:			
parabolic	$X^2 = A+Bt$	$1/X$	1
general parabolic	$X^2+CX = A+Bt$	$2/(2X+C)$	$2X/(2X+C)$
cubic	$X^3 = A+Bt$	$2/X$	2
quartic	$X^4 = A+Bt$	$3/X$	3
logarithmic	$X = A+B\ln(t+t_0)$	$1/B$	X/B

VOL. 14 1960

probably less than 10^{-5}, consistent with the results of HANEMAN's[8] electron diffraction studies on cleaned InSb surfaces.

Nevertheless, the sorption of a monolayer of oxygen is readily attained at 26°C on InAs, GaAs and InP, while two or more monolayer equivalents are rapidly sorbed by the antimonides. In each case the sorption of oxygen continues at an ever diminishing rate, and the surface apparently approaches a metastable configuration. The rate of approach to this configuration decreases, for a given B^V atom, in the order Al>In>Ga.

The kinetic characteristics associated with these configurations are best examined through using a simple set of empirical functions which arise from a general consideration of protective surface reactions.

The kinetic analysis of protective surface reactions

A "protective" surface reaction is any heterogeneous process in which the rate decreases although the external variables are manifestly constant. Examples are concentration polarization, chemisorption, catalytic poisoning, tarnishing, scale formation and indiffusion. The effect generally originates in a chemical or electrochemical change in the configuration of the interface across which the reaction proceeds. The fundamental variable is, therefore, the extent of reaction (X) rather than the time of reaction (t). If $V = dX/dt$ is the rate for given values of the external variables, it is generally more meaningful to examine the empirical function $V(X)$ than $V(t)$. An obvious example lies in the estimation of activation energies, where it is only the temperature dependence of $V(X)$ that has significance.

When "protection" results from the creation of a new phase at the interface, as in chemisorption (where X is the fractional coverage of the surface), or in tarnishing or scaling reactions (where X is the film thickness), two useful functions for data analysis are found to be:

$$\eta = -\partial \log V/\partial X$$

$$\epsilon = -\partial \log V/\partial \log X = \eta X$$

Both functions are always positive for protective reactions, since the rates decrease with X. Their dependence on X for several well-known rate equations are shown in Table 1. The three principal reasons for their utility are:

(1) They permit a classification of the kinetics at any stage of a protective reaction, particularly since their values are in most cases, simple and independent of integration constants.

(2) They readily delineate transitions between rate-controlling reaction steps in the same mechanism or transitions between different mechanisms.

(3) They are independent of the values of external variables, such as temperature and pressure, if the effects of these variables upon the rate are separable by factoring. They provide, therefore, a useful test for the validity of factoring.

The use of the ϵ-function in delineating a transition between mechanisms in the high temperature oxidation of InSb, and in identifying the nature of these mechanisms was illustrated in the earlier

Table 2. Composition of the metastable configurations

Compound	N_{max} (26°)	N_{max} (71°)	N^*_{max}
	oxygen atoms/cm^2 × 10^{-15}		oxygen atoms per surface atom
InSb	$2{\cdot}0_0$	$2{\cdot}4_6$	$>3{\cdot}8$
GaSb	$2{\cdot}4_6$	$2{\cdot}4_1$	$3{\cdot}2_4$
AlSb	$2{\cdot}6_5$	$2{\cdot}7_2$	$3{\cdot}5_1$
InAs	$1{\cdot}4_2$	$1{\cdot}4_0$	$1{\cdot}8_9$
GaAs	$1{\cdot}3_0$	$1{\cdot}2_6$	$1{\cdot}4_6$
InP	(>1·05)		(>1·3)

papers of this series. The present results will now be considered in terms of the η-function.

The transition from "chemisorption" to oxide formation in the $A^{III}B^{V}$ *compounds*

In the present experiments the variable N is a measure of the extent of reaction. Thus N is identified with X and dN/dt with V in the general expression of η and ϵ. In the present context, η is the more useful function and can be derived from the data of Fig. 1. The results are plotted against N in Fig. 2.

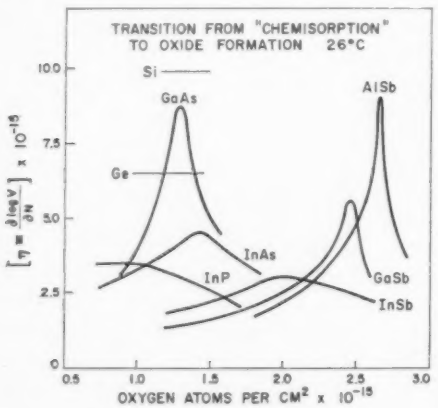


FIG. 2. Transitions from "chemisorption" to oxide formations.

Except for InP,† η initially rises with increasing N. This is characteristic of an approach to a stable configuration in which η would be infinite. The configuration is never attained, however, since the compounds are thermodynamically unstable in oxygen. The beginnings of true oxide formation, presumably involving thermal or field induced diffusion of crystal defects, are clearly seen through the subsequent fall of η. A natural consequence of the latter mechanism is the approach of ϵ to a constant value between 0·5 and 2·0 as N increases.[2, 9] This has been verified for InSb at 212°C[10] where ϵ becomes equal to 1·4 at $N = 5 \times 10^{15}$ atoms/cm^2 and remains unchanged to $N = 11 \times 10^{15}$ atoms/cm^2.[10]

It seems probable that, in each case, a value of N exists which uniquely defines the metastable configuration. While the proper value is obscured by the onset of the next stage of reaction it is likely to be close to N_{max}, where the maximum in η is observed. The magnitude attained by η in the region of N_{max} is a measure of the stability of the configuration; the stability decreases, accordingly, in the order AlSb≈GaAs>GaSb>InAs>InSb.

If N_{max} does define a unique system it should be invariant with temperature. Preliminary experiments at 71°[10] give values of N_{max} which are virtually the same as at 26° for AlSb, GaSb, InAs and GaAs; the value for InSb is shifted, however, to a significantly higher value (Table 2).

† InP differs from the remaining compounds in that its initial reaction rate is logarithmic; i.e. η = constant. The existence of a distinct metastable system is thus questionable. In this respect the oxidation of InP at 26°C resembles that of Ge at 360°C where η is initially constant but then falls off, leading to a constant value of ϵ.[5]

Table 3. Estimation of surface atom density of crushed $A^{III}B^{V}$ *compounds, germanium and silicon*

Material	Estimated[12] proportions on fractured surfaces			a_0(Å)	(atom/cm²)
	(111)	(110)	(100)		
InSb	0·25–0·30	0·70–0·75	0	6·48	$6{\cdot}45\times10^{-14}$
GaSb	0·20–0·25	0·75–0·80	0	6·09	$7{\cdot}35\times10^{-14}$
AlSb	0·05–0·10	0·90–0·95	0	6·10	$7{\cdot}60\times10^{-14}$
InAs	0·15–0·20	0·80–0·85	0	6·06	$7{\cdot}45\times10^{-14}$
GaAs	0·08–0·15	0·85–0·92	0	5·63	$8{\cdot}78\times10^{-14}$
InP	0·15–0·20	0·80–0·85	0	5·86	$8{\cdot}00\times10^{-14}$
Ge	0·65–0·70	0·05	0·25–0·30	5·65	$7{\cdot}00\times10^{-14}$
Si	0·60–0·70	0·05	0·30–0·35	5·43	$7{\cdot}40\times10^{-14}$

Surface atom densities:

$$(111) = 2/(\sqrt{2}a_0^2),$$
$$(110) = 4/(\sqrt{3}a_0^2),$$
$$(100) = 2/a_0^2$$

The surfaces obtained by fracturing the $A^{III}B^{V}$ compounds lie normal to some direction in the ⟨110⟩–⟨111⟩ crystallographic zone. Any such plane may be regarded as a stepped mixture of pure (110) and (111) planes. Since the proportion of the two has been estimated by WOLFF[11], it is possible to calculate the average atom density on surfaces exposed by crushing. The appropriate factors for converting N to N^*, where N^* is the number of oxygen atoms per surface atom are developed in Table 3. An uncertainty of 10 per cent in the proportions of (110) and (111) character will not significantly affect the results. These factors have been used to convert N_{max} to N^*_{max} in Table 2.

Nature of the intermediate configurations

The most prominent results pertaining to the metastable configurations are (1) the high values of N^*_{max}, particularly for the antimonides, and (2) the distinct grouping of N_{max} for the antimonides and for the arsenides (Table 2). The irreversible quality of the reactions suggests that oxygen atoms, rather than molecules, are ultimately involved in the chemical bonding. The value of 3–4 oxygen atoms per surface atom is too high to be explained plausibly by the chemisorption of oxygen on the original surface atoms alone, so that additional atoms probably take part in the formation of the intermediate structures. One moves, therefore, from the usual concept of chemisorption, where the structure of the adsorbed layer is determined primarily by the structure of the underlying surface, to the concept of complex formation, in which the film can have a chemical and structural identity of its own.

In this connection it is significant that the distinction between the arsenides and the antimonides (Table 2) has a parallel in the classical oxygen chemistry of the Group V elements.[12] While antimony is octahedrally coordinated by oxygen, arsenic and phosphorous are tetrahedrally coordinated in their oxygen complexes. No such distinction exists among the Group III elements where octahedral coordination is generally observed. Thus, a large shift in N_{max} in going from the arsenides to the antimonides, and a small effect of changing the A^{III} atoms, is more consistent with the behavior of the atoms in their normal oxygen compounds than with the behavior to be expected of atoms that are still covalently attached to the surface of a zincblende structure.

The conclusion to be drawn from this admittedly heuristic argument is that the composition and structure of the complex depends more upon the inherent chemical properties of the atoms in the substrate rather than on the properties peculiar to the substrate itself. It should be noted that one cannot, on the basis of the present data, exclude the possibility that complex formation proceeds through several stages in which several discrete

structures may be involved. Further work is necessary before one can define the structure of the complexes and the mechanisms by which they are formed and superseded.

Acknowledgement—The author wishes to acknowledge the expert technical assistance of Mr. A. A. MENNA.

REFERENCES

1. ROSENBERG A. J. and LAVINE M. C., submitted to *J. Electrochem. Soc.*
2. ROSENBERG A. J., submitted to *J. Electrochem. Soc.*
3. GREEN M., KAFALAS J. A. and ROBINSON P. H., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 349. University of Pennsylvania Press, Philadelphia (1957).
4. ROSENBERG A. J., ROBINSON P. H. and GATOS H. C., *J. Appl. Phys.* **29,** 771 (1958).
5. ROSENBERG A. J., to be published.
6. ROSENBERG A. J., *J. Amer. Chem. Soc.* **78,** 2429 (1956).
7. BRUNAUER S., EMMETT P. H. and TELLER E., *J. Amer. Chem. Soc.* **60,** 309 (1938).
8. HANEMAN D., this Conference, *J. Phys. Chem. Solids* **14,** 162 (1960).
9. GRIMLEY T. B., *Chemistry of the Solid State* (edited by GAINER W. E.) p. 336. Butterworths, London (1955).
10. ROSENBERG A. J., to be published.
11. WOLFF G. A., private communication.
12. WELLS A. F., *Structural Inorganic Chemistry*, p. 481. Oxford University Press (1952).

VOL
14
196

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 181–185. Printed in Great Britain.

SECTION V. ORDINARY SURFACES OF GERMANIUM AND SILICON

OBSERVATIONS OF SURFACE TRANSVERSE MAGNETORESISTANCE EFFECTS ON *n*-TYPE GERMANIUM*

W. A. ALBERS, Jr.† and J. E. THOMAS, Jr.‡
Wayne State University, Detroit, Michigan

VOL. 14 1960

Abstract—New techniques of preparing thin semiconductor samples of uniform thickness have been developed, which allow for high sensitivity in the measurement of surface effects. Measurements of the surface transverse magnetoresistance effect first reported by ZEMEL and PETRITZ have been extended to lower resistivity *n*-type material of different crystallographic orientation. Data were obtained for directions of the magnetic field both in and perpendicular to the plane of the thin films. Large swings in surface potential were accomplished by employing an HF rinse prior to ambient cycling. Observations of surface transverse magnetoresistance effects as a function of surface potential at fixed magnetic field strengths, and as a function of magnetic field strength at fixed surface potentials are reported. The analysis of the data requires a general solution of Poisson's equation which does not employ the semi-infinite geometry in order to ascertain values of the electric field at the surface of a semiconductor. This general solution is presented for several special cases of sample thickness and resistivity and the deviations from the semi-infinite geometry solutions are delineated. Finally, the observations of the surface transverse magnetoresistance effect are discussed in view of the present theoretical model of the surface.

1. INTRODUCTION

STUDIES of the surface properties of semiconductors through observations of surface transport phenomena require a knowledge of the effect of the surface scattering of carriers which is reflected in the carrier mobilities. SCHRIEFFER[1] has developed a theory of the reduction of the conductivity mobility for carriers near the surface of a semiconductor subject to the condition of random scattering at the surface. Recently, ZEMEL[2] extended this theory to include transverse magnetic fields. Using expressions developed by PETRITZ[3] for the surface galvanomagnetic coefficients as functions of effective mobilities and surface carrier concentrations, ZEMEL and PETRITZ[4, 5] were first to experimentally study surface Hall and magnetoresistance coefficients for shallow space-charge potential wells.

MISSMAN and HANDLER[6] have recently reported observations of the surface Hall coefficient for deep space-charge potential wells on *n*-type germanium. These studies have served qualitatively to verify the reduction in mobilities due to random surface scattering of carriers and, further, have pointed out the important role of light holes in surface transport phenomena.

We have obtained simultaneous measurements of conductivity and magnetoresistance on relatively low resistivity *n*-type germanium employing special surface treatments in order to extend the

* This work was supported by Air Research and Development Command, Department of the Air Force, AF 49 (638)–158. Submitted in part by W. A. ALBERS, Jr. to the Graduate School of Wayne State University as partial fulfilment for the requirements for the Doctor of Philosophy Degree in physics.

† Present address: Bendix Aviation Corporation, Research Laboratories Division, Detroit, Michigan.

‡ Present address: Sylvania Electronic Products, Inc. Semiconductor Division, Woburn, Mass.

data to space-charge potential well depths on *n*-type material intermediate between those reported by ZEMEL and PETRITZ[4] and MISSMAN and HANDLER[6].* Large sample surface-to-volume ratios were accomplished by employing recently developed etching techniques, resulting in uniform sample thicknesses of the order of 1–10 μ (10^{-4}–10^{-3} cm). We have further observed the magnetic field strength dependence of the surface magnetoresistance effect for small magnetic fields at constant space-charge potential well configurations.

2. EXPERIMENTAL

Sample blanks were prepared in the form of rectangular slabs cut from oriented single crystal germanium boules. The blanks were subsequently shaped to the form shown in Fig. 1 by using a

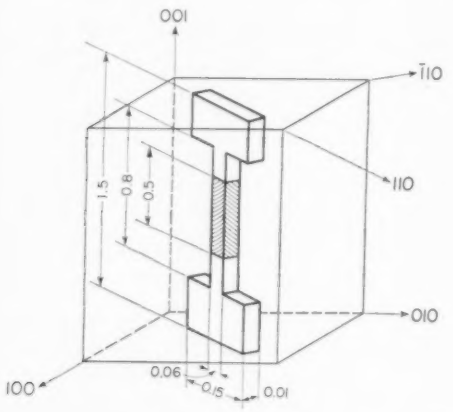


FIG. 1. Details of sample dimensions and orientation for magnetoresistance measurements. All dimensions are in centimeters.

White "Airbrasive" sandblasting unit. Sample orientation and approximate dimensions are also delineated in Fig. 1. After lapping to a suitable thickness and soldering contacts at the ends, the samples were mounted on glass forms for support. A rectangular hole was centered in the glass form so that the filament portion of the sample was free.

The thin film region of the sample is indicated by the shaded area in Fig. 1. This was accomplished by two different etching techniques. The first was a photoelectrolytic etching method described by ALBERS and THOMAS[7], wherein the *n*-type semiconductor is made the anode of an electrolytic cell and is illuminated with penetrating radiation in the area to be etched. The second approach consisted of a stream chemical etch using the H_2O_2:HF:H_2O system. Control of final thickness was obtained by water dilution of the etchant as etching progressed. The etchant stream was constrained to the desired area to be etched by polyethylene guide strips. In both cases, the resulting sample geometry was such that the thin samples had bulky end contacts, and the total resistances of the samples were consequently dominated by the resistance of the thin portions. In this manner, several *n*-type germanium samples were prepared which had uniform thicknesses in the range of 2–10 μ, as measured by the transmission of a beam of X-radiation collimated by a 0·001-in. diameter pinhole.

The method of measuring conductivity and magnetoresistance effects followed that reported by ZEMEL and PETRITZ[4], as depicted schematically in Fig. 2. The d.c. amplifier used was the General Radio 1230–A, the output of which was fed, through a potential divider network, to a recording potentiometer. With no applied magnetic field, the conductivity voltage across the sample was bucked out.

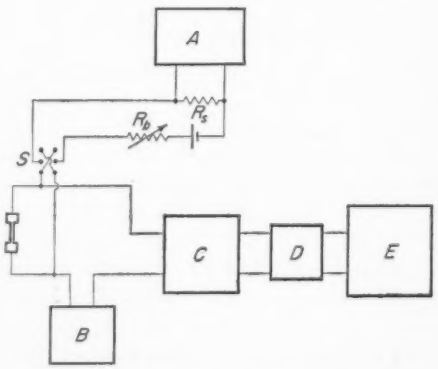


FIG. 2. Schematic representation of measuring apparatus.

A. Leeds Northrup type-K2 potentiometer.
B. Bucking voltage source.
C. D.C. amplifier and electrometer.
D. Recorder sensitivity control.
E. Leeds Northrup Speedomax recorder.
R_b. Sample ballast resistor.
R_s. Standard resistance.
S. Current reversing switch.

* This range of surface potentials has been covered on *p*-type intrinsic material.[5]

VOL. 14 1960

Upon placing the sample in the field of a 4-in. electromagnet, the resulting change in sample resistance was observed directly on the recorder.

The surface potential and, consequently, the shape and depth of the space-charge potential well, were varied by exposing the sample to atmospheres of wet nitrogen, dry oxygen and dry ozone. The use of a gas proportioner allowed for various mixtures of these gases to be used as ambient, facilitating measurement at constant surface potential. By rinsing the sample in HF prior to placing in the sample holder, relatively deep potential wells appeared on n-type germanium corresponding to highly p-type surfaces. This was checked by small signal a.c. field effect observations and also was consistent with the observed changes of conductivity with changing ambient atmosphere. As the HF-treated surface aged, the surface conductivity drifted back in the direction of the n-type condition. Upon ambient cycling at successive intervals during this drift, a wide range of surface type was obtained. A tentative explanation of the results of the HF rinse is that the HF removes some of the oxide layer, apparently to the extent that the surface begins to approach the "clean" state described in Ref. (6), resulting in a surface which is p-type. Subsequent regrowth of the oxide then is associated with the drift towards an n-type surface. On the other hand, the purity of the HF used for the rinse was in poor control due to the presence of organic masking material that protected sample leads, and it could very well be that impurities were contributing to the observed effects.

3. RESULTS AND DISCUSSION

Measurements were performed on several n-type samples in the resistivity range of 7–12 Ω-cm. The results were qualitatively similar for all of the samples in this range and we thus restrict our attention to a representative 8·5 Ω-cm sample. There were, further, no apparent differences in samples prepared by the photoelectrolytic and chemical stream etching methods.

Measurements of the transverse magnetoresistance perpendicular to the sample surface vs. reduced conductivity change are plotted in Fig. 3. The points were obtained by smoothing data which were obtained from several runs on several different days. The solid curve in Fig. 3 results from best fit calculations using the expressions of PETRITZ[3] for the three-band model, neglecting correlation and bulk hole terms. The best fit parameter was the ratio of light to heavy holes, resulting in a value for this ratio of 0·022, in good agreement with other workers.[8, 11] The values of effective

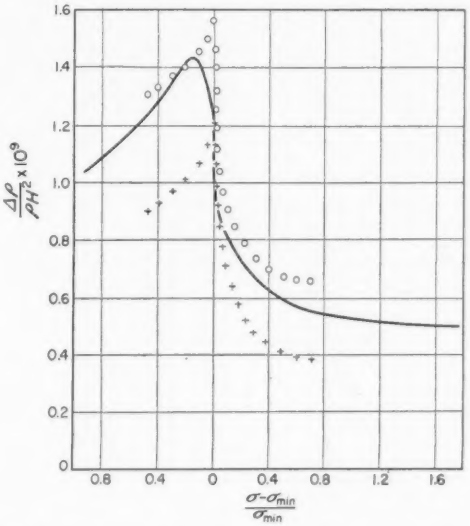


FIG. 3. Experimental and theoretical magnetoresistance coefficients for 8·5 Ω-cm n-type germanium (sample M-8): H in (110) directions; I in (100) direction; sample thickness of 7×10^{-4} cm.

○ Experimental
— Theoretical
} $H_\perp$ + Experimental $H_\parallel$

mobilities used in the calculations were obtained from the work of ZEMEL and SCHRIEFFER[2] for the linear space charge model. The magnetoresistance effective mobility was assumed to be equal to the Hall effective mobility.

The values of excess carriers at the surface used in the calculations were determined by a general solution of Poisson's equation. Previous solutions of this equation have been based on a semi-infinite geometry,[8–10] which no longer can be expected to apply for thin samples. It was found that the semi-infinite geometry solutions apply for sample thicknesses greater than αL_D, where α is a function of the initial bulk resistivity. The value of α must be determined for each resistivity by numerical integration of Poisson's equation. (A plot of α vs. ρ is presented in Fig. 4 for germanium as obtained

from numerical solutions of Poisson's equation.) The appropriate general solutions for the surface excesses were used for films that were thinner than the corresponding value of αL_D.

The choice of the reduced conductivity change as abscissa in Fig. 3 was made for ease of comparison with the data of ZEMEL and PETRITZ[4, 5], The incorporation of these factors in the development of analytical expressions for the galvanomagnetic coefficients is expected to yield a more realistic comparison of experiment and theory.

The range of surface potential represented by the data of Fig. 3 can best be appreciated with the aid of Fig. 5. Here is plotted the theoretical curve

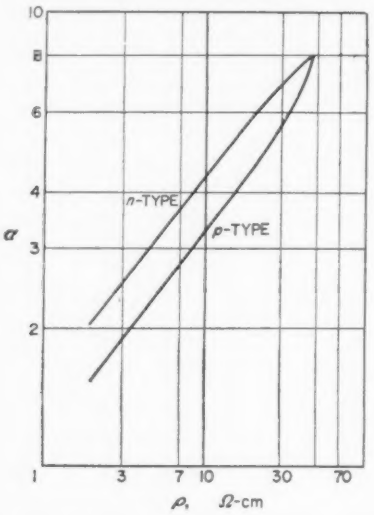


FIG. 4. Plot of α vs. resistivity, ρ, defining germanium sample thicknesses for which the surface excesses do not deviate more than 1 per cent from those obtained in Ref. (8) for a semi-infinite geometry.

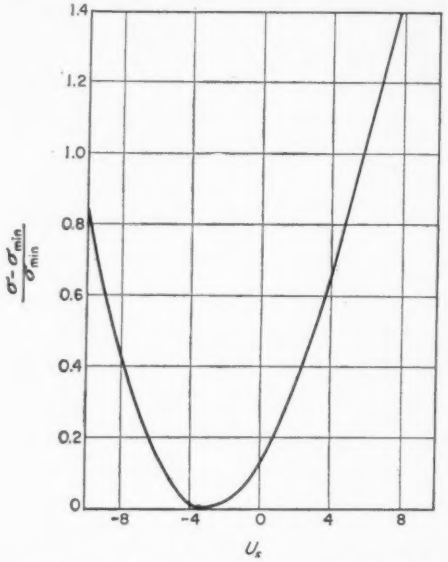


FIG. 5. Reduced conductivity change vs. reduced surface potential for 8·5 Ω-cm n-type germanium calculated from theory.

and the shape of the present curve is qualitatively similar to their published results. Of particular interest is the striking peak in the experimental points near the conductivity minimum which is attributed to the light hole contributions. The experimental and theoretical curves are in fair qualitative agreement, and the discrepancies might be due to a number of factors:

(a) the energy dependence of the relaxation time was neglected in the development of the theoretical curves;

(b) the linear space charge approximation is inaccurate;

(c) the many-valley nature of the conduction band was neglected;

(d) light holes were assumed to obey the same effective mobility variations as heavy holes.

of the reduced conductivity as a function of reduced surface potential $u_s = q\phi_s/kT$. Upon comparison with Fig. 3, it will be noted that the surface potential ranges from approximately $+5$ to -9 units of kT, the -9 corresponding to a highly p-type surface on the n-type germanium sample. Comparing these results with those of Refs. (4) and (6) indicates that our range of surface potential bridges the gap in the p-type surface potential that existed in previous work.

Also plotted in Fig. 3 are the smoothed experimental data obtained for the transverse magnetoresistance coefficient with the field in the plane of the thin sample. Analytical expressions developed parallel to PETRITZ'[3] approach assuming a constant Hall field in the direction of the thin dimension of the sample failed to predict the observed differences between the perpendicular and parallel

field cases. The assumption of a constant Hall field is undoubtedly poor here, and further work on the problem of crossed electric and magnetic fields in the plane of thin semiconductor will be necessary in order to analyse the data satisfactorily.

The low magnetic field dependence of the transverse magnetoresistance effects are plotted in Fig. 6 for several different constant values of surface holes in surface transport phenomena. The gap in potential well depth in the reported work of others[4, 6] on *n*-type germanium has been bridged and the reduction in mobility is qualitatively similar. There is no evidence that the low magnetic field surface magnetoresistance effect cannot be interpreted in terms of bulk transport mechanisms, provided surface scattering is taken into account.

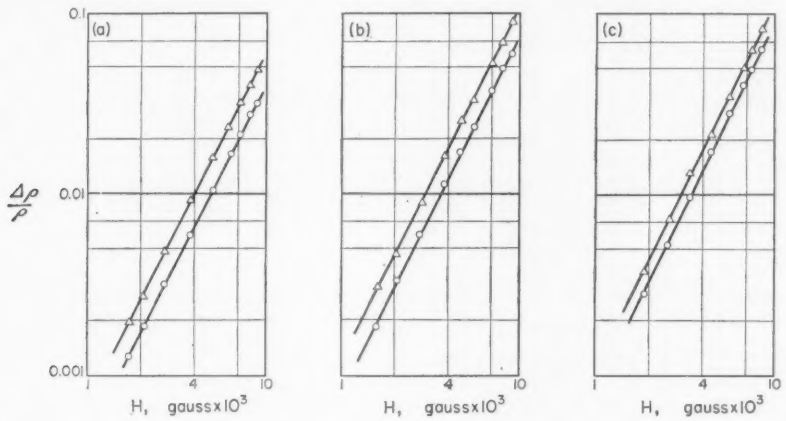


FIG. 6. Magnetoresistance effect vs. magnetic field at several values of constant surface potential: H in (110) directions; I in (100) direction; *n*-type germanium (M-8) thickness of 7×10^{-4} cm. (a) $(\sigma-\sigma_{min})/\sigma_{min} = 0{\cdot}50$ *n*-side; (b) $(\sigma-\sigma_{min})/\sigma_{min} = 0{\cdot}04$ *n*-side; (c) $(\sigma-\sigma_{min})/\sigma_{min} = 0{\cdot}41$ *p*-side.
Δ $H_\perp$ $\bigcirc$ $H_\parallel$

potential. In the range of magnetic field strengths used, the data obey the H^2 dependence characteristic of bulk magnetoresistance effects. It will be noticed that there is no apparent difference in the perpendicular and parallel field cases other than the reduction in the magnitude of the effect for parallel fields. It is concluded that if the transport phenomena at the surface rely on different mechanisms from those in the bulk, they are not observable at the low field strengths investigated. This, of course, does not rule out the possibility that large magnetic field observations may yield indications of deviations from the bulk transport mechanism.

4. CONCLUSIONS

The data presented verifies the importance of the random scattering process and the role of light

REFERENCES

1. SCHRIEFFER J. R., *Phys. Rev.* **97,** 641 (1955).
2. ZEMEL J. N., *Phys. Rev.* **112,** 762 (1958).
3. PETRITZ R. L., *Phys. Rev.* **110,** 1254 (1958).
4. ZEMEL J. N. and PETRITZ R. L., *Phys. Rev.* **110,** 1263 (1958).
5. ZEMEL J. N. and PETRITZ R. L., *J. Phys. Chem. Solids* **8,** 102 (1959).
6. MISSMAN R. and HANDLER P., *J. Phys. Chem. Solids* **8,** 109 (1959).
7. ALBERS W. A. Jr. and THOMAS J. E. Jr., *Bull. Amer. Phys. Soc.* **3,** 219 (1958).
8. KINGSTON R. H. and NEUSTADTER S. F., *J. Appl. Phys.* **26,** 718 (1955).
9. GARRETT C. G. B. and BRATTAIN W. H., *Phys. Rev.* **99,** 376 (1955).
10. DOUSMANIS G. C. and DUNCAN R. C. Jr., *J. Appl. Phys.* **29,** 1627 (1958).
11. WILLARDSON R. K., HARMAN T. C. and BEER A. C., *Phys. Rev.* **96,** 1512 (1954).

VOL. 14 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 186–192. Printed in Great Britain.

SURFACE MOBILITY MEASUREMENTS IN GERMANIUM

A. MANY, N. B. GROVER, Y. GOLDSTEIN and E. HARNIK

Department of Physics, The Hebrew University, Jerusalem, Israel

Abstract—Measurements are reported of the surface mobility in accumulation layers of germanium at low temperatures, and results obtained are compared with the theoretical calculations of SCHRIEFFER. Advantage is taken here of the relatively slow thermal release of charge from many of the surface states into the majority-carrier band at sufficiently low temperatures. By the use of the pulsed field-effect technique, a rapid decrease in majority-carrier concentration within the space-charge region is effected. The change in surface conductance measured at the onset of the pulse, before the deficit induced can be redistributed between the states and the majority-carrier band, then becomes a direct measure of the majority-carrier surface mobility.

The results for *n*-type samples show that SCHRIEFFER's corrections for electrons are overestimated. On the assumption that the theoretical values are valid for accumulation layers, it is concluded that the surface is not a completely diffuse scatterer.

INTRODUCTION

SURFACE conductivity has been extensively employed in the study of semiconductor surfaces as it is the most convenient and accurate measure of the barrier height V_s. The value of V_s can be determined[1–3] from a knowledge of ΔP and ΔN, the surface hole and electron densities in excess of the densities at flat-band conditions ($V_s = 0$). These latter are in turn connected with the change in surface conductivity $\Delta\sigma$ by the equation

$$\Delta\sigma = q(\mu_{pS}\Delta P + \mu_{nS}\Delta N) \qquad (1)$$

where μ_{pS} and μ_{nS} are the effective surface mobilities associated with ΔP and ΔN.

The surface mobilities are expected to be lower than the respective bulk values due to the scattering present at the surface in addition to the normal bulk scattering processes.[3] For small values of $|V_s|$ it is common practice to use the bulk mobility values in equation (1). For strong inversion or deep accumulation layers, however, either the holes or the electrons are constrained to move along a deep potential well and, as a result, the reduction in mobility may become very considerable. Hence within this range a knowledge of the dependence of the surface mobility on V_s is essential for deducing V_s from measurements of $\Delta\sigma$. More important still, such information should shed light on the nature of the scattering processes at the surface.

Assuming the surface to be a completely diffuse scatterer, SCHRIEFFER[3] has calculated the surface mobilities as a function of barrier height. The nature of his approximations makes the results adequate only for strong inversion layers. Several attempts[4–7] have been made to improve these calculations. In particular, FRANKL[7, 8] has been able to extend the application of the theory to small $|V_s|$ values and to accumulation layers.

It would be highly desirable to check experimentally whether the diffuse-scattering model is valid. Such measurements are rendered difficult, however, by the presence of fast surface states. Were these states absent, the surface mobility would have been easily deducible from field-effect measurements. In this case ΔN and ΔP would be given directly by the known induced charge.[1] But, because of the surface states present, the fraction of the induced charge which is mobile is not known, and the measured change in surface conductivity yields only the product $\mu_{pS}\Delta P$ or $\mu_{nS}\Delta N$. Additional information is thus necessary to separate the two unknowns. One possibility is to measure the space-charge capacitance,[9, 10] and from this to evaluate ΔP or ΔN. Such measurements are difficult as the series geometric capacitance cannot be made large enough to permit accurate detection of changes in space-charge capacitance which, in the region of interest, is of the order of $0{\cdot}1\,\mu F/cm^2$. Another possibility,

VOL. 14 1960

employed by PETRITZ and ZEMEL[11–14], consists of measuring the surface Hall coefficient and magnetoresistance. These authors conclude that the surface is essentially a diffuse scatterer as assumed by SCHRIEFFER.

A different approach to the problem would be to attempt to reduce the density of the fast states on the surface by suitable chemical treatments. In this manner, most of the induced charge in the field-effect experiment would appear in the space-charge region and would thus contribute to the change in surface conductivity. Measurements along these lines were carried out by MILLEA and HALL[15] on low resistivity samples, and their results indicate that the reduction in mobility predicted by SCHRIEFFER is overestimated.

In this paper the experimental procedure is chosen with the purpose of eliminating the effect of the surface states rather than reducing their density. Advantage is taken here of the relatively slow thermal release of charge from the states into the majority-carrier band at low temperatures. By the use of the pulsed field-effect technique, a rapid decrease in majority-carrier concentration within the space-charge region is effected. The change in $\Delta\sigma$ is then measured at the onset of the pulse, before the deficit induced can be redistributed between the band and the states. In this manner $\Delta\sigma$ can be obtained as a function of the majority-carrier excess density, and from this the surface mobility can be evaluated as a function of V_s.

Results of surface mobility in accumulation layers are presented for several germanium samples. The results obtained are compared with theoretical values based on SCHRIEFFER's calculations.[3] It is found that for deep accumulation layers in n-type samples the experimental points usually lie well above these calculated values. On the assumption that the theoretical data are valid in this range, it is concluded that the surface is not a completely diffuse scatterer. For p-type samples, no definite conclusions can be drawn from the data obtained so far.

EXPERIMENTAL METHOD

The germanium samples studied were rectangular filaments about 10 mm long, 3 mm wide and 0·3 mm thick. Following a chemical etch they were mounted on a jig with thin (10 μ) Mylar spacers separating the two larger surfaces from the field electrodes. The jig was inserted into a cryostat in which the temperature could be controlled between 100 and 300°K. The measurements were carried out in vacuum except those at 80°K which were made while the jig was immersed directly in liquid air.

The purpose of the measurement is to obtain the change in $\Delta\sigma$ at the onset of the pulsed field as a function of pulse amplitude. Instead of varying the pulse amplitude, however, a constant-amplitude pulse δV is used which can be switched on and off the field plate in series with a variable biasing a.c. or d.c. voltage, as shown in Fig. 1(a). The same pulse independently activates the Wheatstone bridge consisting of the filament, a decade resistor, and two fixed resistors. The CRO serves as a null indicator and also to display the changing filament conductance throughout the duration of the pulse. The measuring sequence is as follows. The bridge is balanced and the filament resistance R_f read on the decade resistor. The pulse δV, of polarity such as to repel majority-carriers from the surface, is then switched on at the field plate. The decade resistor is readjusted to balance the bridge as close to the onset of the pulse as the circuitry permits ($\sim$10 μsec.). The pulse is then switched off the plate and the plate voltage (a.c. or d.c.) is adjusted so as to restore bridge balance. The pulse is now switched on again and the process repeated. Figure 1(b) illustrates this sequence for the case where the biasing voltage is a 50-cycle alternating source. The pulse is synchronized[16] with the crest of the a.c. voltage. In this manner a series of values of filament resistance $(R_f)_n$ is obtained, each corresponding to the number of times n that the pulse has been switched on. This is equivalent to a measurement of the filament resistance as a function of $n\delta V$, the voltage of a pulse whose amplitude is varied in steps of δV (with no bias-voltage).

The maximum resistance of the filament, corresponding to the minimum of surface conductance $\Delta\sigma_{\min}$, is derived from the usual d.c. field-effect. The various $(\Delta\sigma)_n$ values corresponding to $(R_f)_n$ are then computed by reference to $\Delta\sigma_{\min}$. If one assumes, as pointed out above, that no surface states are involved in this process, then the mobile charge induced is proportional to $n\delta V$. This enables the evaluation of the barrier height V_s for each value of n.

RESULTS

In the measurements described minority-carrier effects are completely absent, so that the samples studied constitute in effect single-carrier systems. This is insured by using extrinsic material at low temperatures and confining the equilibrium value of the barrier height to the appropriate region.

In Fig. 2 are shown the results of measurements on an n-type sample at 80°K. In Fig. 2(a) $\Delta\sigma$ is plotted as a function of the effective pulse amplitude $n\delta V$ ($\delta V = -24$ V) applied to the field plate as described in the previous section. As all the induced charge is assumed to appear as a change in the concentration of mobile majority-carriers

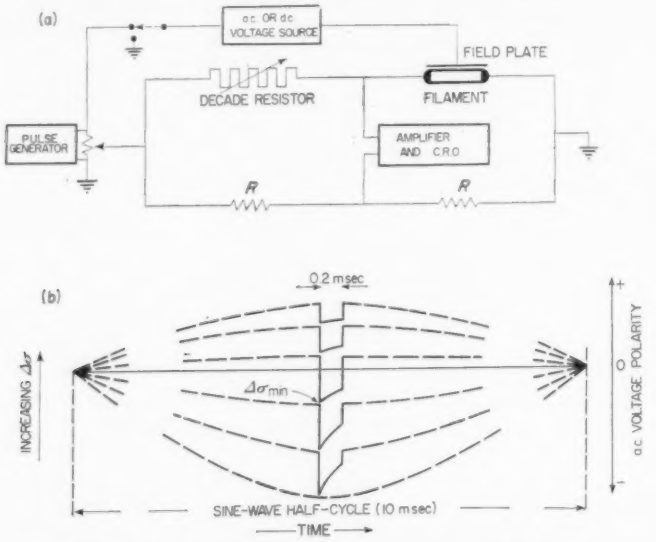


FIG. 1. (a) Schematic outline of the experimental set-up. (b) Illustration of the measuring sequence for deriving $\Delta\sigma$ as a function of $n\delta V$. The heavy solid lines correspond to the display on the CRO.

Only when the minimum of surface conductivity is being determined do minority carriers play any role. In this latter procedure time is allowed for the surface states and the two bands to attain equilibrium conditions. Actually, however, even at the minimum the surface excess of minority carriers is negligible compared to the surface excess (or rather deficit) of majority carriers.

Once $\Delta\sigma_{\min}$ has been determined, the measurements are made under conditions where the majority-carrier band is in equilibrium with the induced charge whereas the surface states and minority-carrier band are not. These conditions exist for a.c. and pulsed fields at sufficiently low temperatures. In such measurements it is possible to swing the majority-carrier band over a considerable range in the depletion region, even far beyond the position corresponding to $\Delta\sigma_{\min}$.

(electrons), the barrier height V_s can be computed at any point along the curve from the experimentally-known induced charge and from the theoretically-known[1, 2] value of ΔN corresponding to $\Delta\sigma_{\min}$. The initial bias voltage is so chosen in the figure that the origin corresponds to flat-band conditions ($V_s = 0$). Consider first the negative quadrant ($V_s < 0$). It is seen that the experimental points lie on a straight line up to $V \cong -80$ V. This corresponds to values of $|V_s|$ up to several times the forbidden gap (Fig. 2b). For larger $|V|$ the experimental points (circles) begin to deviate from the straight line due to the progressively decreasing space-charge capacitance as $|V_s|$ increases. The induced charge $q\Delta N$ is no longer proportional to V but to $V - V_s$. The crosses in Fig. 2(a) are replots of $\Delta\sigma$ versus $V - V_s$. These are seen to fall exceedingly well on the extension of the straight

line obtained for smaller values of $|V_s|$. The slope of this straight line is in remarkable agreement with that computed from the electron bulk mobility. Besides the obvious conclusion that the surface mobility is equal to the bulk mobility for large negative barriers, these results prove conclusively that in this region there are no surface states characterized by release times smaller than the resolution time of the experimental apparatus ($\sim$10 μsec). It should be noted moreover that it was possible to elevate the conduction band up to about 60 times the forbidden gap value (Fig. 2b) without the measurements at the onset of the pulse being disturbed by minority-carrier effects.

The surface mobility begins to decrease as the barrier moves toward accumulation-layer conditions (positive V and V_s). This region, of principal interest in the present work, is shown in greater detail in Fig. 3(a). Here $\Delta\sigma$ is plotted as before but the abscissa is marked with the corresponding values of barrier height v_s measured in units of kT/q. For any particular value of v_s the ordinate of the experimental curve is equal to $q\Delta N\mu_S$. The corresponding value of $q\Delta N\mu_B$ is the ordinate of the point in question on the extension of the straight line that represents $\Delta\sigma$ for $V_s<0$. This value is just what $\Delta\sigma$ would be had the carriers continued to move with their bulk mobility

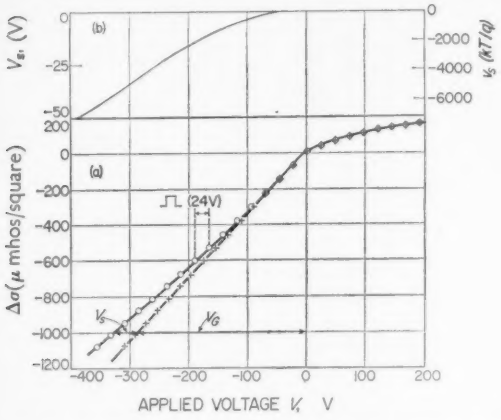


FIG. 2. (a) Typical results (circles) of $\Delta\sigma$ versus the effective pulse-voltage V_i for an n-type sample ($T = 80°$K). The crosses represent a replot of $\Delta\sigma$ versus $V-V_s$, where V_s is the barrier height. (V_G is the voltage drop across the geometric plate-to-surface capacitance C_G.)

$$V = V_G + V_s$$

$$q\Delta N = C_G V_G = C_G(V - V_s)$$

$$\Delta N \simeq \text{const. } X\sqrt{V_s}$$

(b) V_s versus effective pulse-voltage, as computed from the equations appearing above.

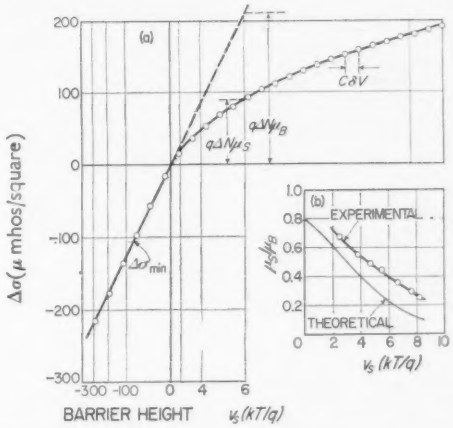


FIG. 3. (a) The same as Fig. 2(a), but with the accumulation region in greater detail. Here $C\delta V$ is the induced charge corresponding to the pulse used. Values of the barrier height v_s in units of kT/q are shown. (b) A plot of the experimental values of μ_S/μ_B versus v_s, together with the corresponding theoretical curve.

μ_B. Thus μ_S/μ_B at any value of v_s is just the ratio of the two ordinates at that point.

The ratio of surface to bulk mobility, as computed from the experimental curve of Fig. 3(a), is shown in Fig. 3(b) as a function of v_s. Also shown is the theoretical curve according to SCHRIEFFER[3], slightly corrected as suggested by FRANKL[8]. It is seen that the experimental points are well above SCHRIEFFER's values.

Figure 4 represents results of measurements of μ_{nS}/μ_{nB} as a function of v_s made on an n-type sample at three different temperatures. Also shown are the corresponding theoretical curves according to SCHRIEFFER. It is seen that the theoretical curves in all three cases cross the experimental curves, thus showing that the former are overestimated for deep accumulation layers. The experimental curves are lower, the lower the temperature, as is to be expected from the increasing mean free path of the electrons as the

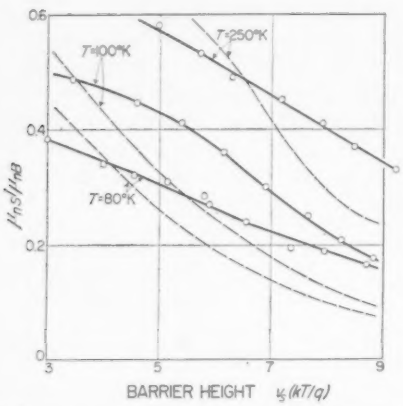


FIG. 4. Experimental results of μ_{nS}/μ_{nB} versus v_s for an n-type sample ($N_D-N_A = 6{\cdot}6\times10^{13}$) at three different temperatures, together with the corresponding theoretical curves.

——— Experimental, — — — Theoretical

temperature is decreased. Quite a different behavior is exhibited in Fig. 5, which represents measurements of μ_{nS}/μ_{nB} versus v_s for another n-type sample. Here the results at the two temperatures 80°K and 152°K almost coincide, whereas the theoretical curves are of course different.

Figure 6 shows surface mobility results for four n-type samples of three different impurity contents.

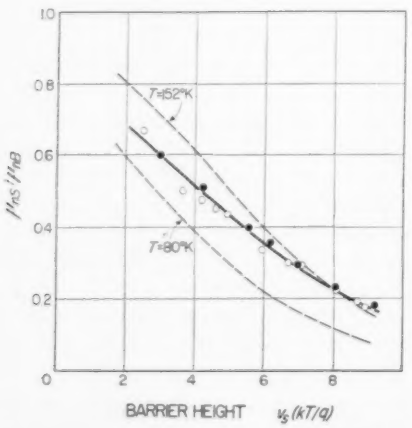


FIG. 5. Experimental results of μ_{nS}/μ_{nB} versus v_s for an n-type sample ($N_D-N_A = 4{\cdot}4\times10^{13}$) at two different temperatures, together with the theoretical curves.

○ = 152°K, ● = 80°K,
——— Experimental, — — — Theoretical

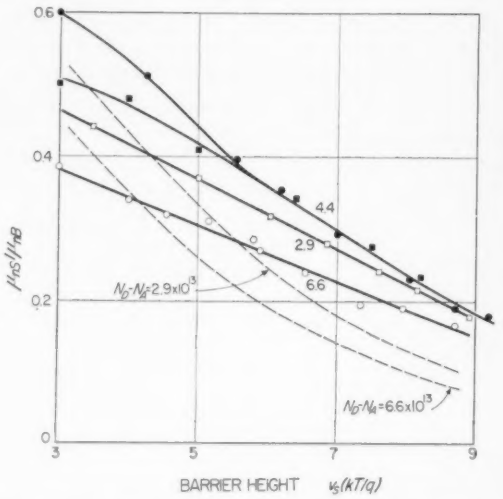


FIG. 6. Experimental results of μ_{nS}/μ_{nB} versus v_s for n-type samples of various impurity content ($T = 80°K$). All the samples were etched in CP–4A, and sample (b) underwent an additional etch in warm H_2O_2.

□ $N_D-N_A = 2{\cdot}9\times10^{13}$
■ $N_D-N_A = 4{\cdot}4\times10^{13}$ (a)
● $N_D-N_A = 4{\cdot}4\times10^{13}$ (b)
○ $N_D-N_A = 6{\cdot}6\times10^{13}$

——— Experimental, — — — Theoretical

One expects that due to the decreased barrier thickness, the higher the impurity content the greater the reduction in mobility. This trend is not fully substantiated, as the curve for the purest sample lies between the other two. The curves for the two different samples of equal resistivities however, coincide at deep accumulation layers. Sample (a) was etched as usual in CP–4A whereas sample (b) underwent an additional etch in warm H_2O_2.

In Fig. 7 are shown results obtained on a p-type sample at two temperatures. The theoretical curves here lie above the experimental curves, but in both cases their separation decreases at deep accumulation layers.

CONCLUSIONS

Results have been presented of surface mobility versus barrier height in accumulation layers of germanium at low temperatures. In the analysis of the data it has been tacitly assumed that the experimental procedure employed eliminated

effects arising from surface states. Actually, the occurrence of any states with release times shorter than the experimental resolution time ($\sim 10\ \mu$sec) would result in the mobile charge being less than the total charge induced by the pulse. The effect of such states, if present, would then be to raise the surface mobility above its measured values on the one hand and, on the other hand, to lower the materially changed. Even in the extreme case where it is assumed that no reduction in mobility takes place over the entire accumulation region, the reduction in the maximum value of v_s is only from 9·5 to 7·2 kT/q. It is thus seen that the experimental curves shown in Figs. 3–7 represent the lowest possible limit to the actual values. For deep accumulation layers in n-type samples, these

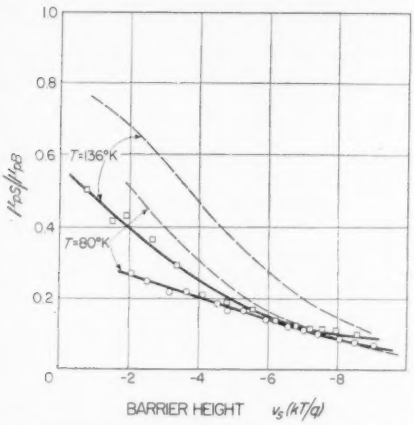


FIG. 7. Experimental results of μ_{pS}/μ_{pB} versus v_s for a p-type sample ($N_A - N_D = 13{\cdot}2 \times 10^{13}$) at two different temperatures.

□ = 136°K, ○ = 80°K,

——— Experimental, — — — Theoretical

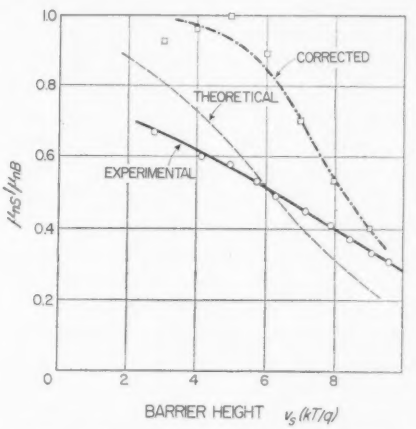


FIG. 8. Experimental results of μ_{nS}/μ_{nB} versus v_s for an n-type sample ($N_D - N_A = 6{\cdot}6 \times 10^{13}$, $T = 250$°K). The dash–dot curve is computed from the experimental data on the assumption that surface states are present with time-constants shorter than the resolution time of the experimental apparatus.

○——○ Experimental, — — — Theoretical,

□ — · — · □ Corrected

values of v_s below those calculated on the basis of the absence of surface states. As v_s for deep accumulation layers depends only logarithmically on the induced mobile charge $q\Delta N$, however, the deduced values of v_s would not be appreciably affected in this range. This argument is illustrated in Fig. 8. The experimental curve represents mobility results on an n-type sample obtained, as in the previous cases, on the assumption that no states are involved in the measurement. The "corrected" curve was constructed on the rather arbitrary assumption that states whose release time is too short to be detected were present with a density of $2 \times 10^{11}/\text{cm}^2$ at 6 kT below the conduction band. (The Fermi level at flat bands lies 12 kT below the conduction-band edge.) The parameters of the states were chosen so as to make μ_{nS} as near as possible to μ_{nB} for small v_s values. It is seen that the "corrected" values lie above the experimental curve but that the v_s values are not lowest limits are seen to be well above the theoretical curves.

The theoretical values used here are based on SCHRIEFFER's calculations which, as pointed out previously, are adequate for strong inversion layers. For lack of other data, the theoretical curves shown were constructed under the presumption[8] that SCHRIEFFER's results are valid for deep accumulation layers as well. It should also be pointed out that, for the samples studied and in the temperature range covered, the mean free path of the carriers is comparable to the thickness of the accumulation layers. This seems to indicate[17] that the calculated reduction in mobility in this case would probably be even greater than that derived by SCHRIEFFER. Thus if the theoretical curves shown in the figures are taken to represent

the correct values based on a completely-diffuse scattering model, it follows from the experimental results presented here that the surface has to be at least partially specular. A definite conclusion on this issue needs await the accurate computations which are reportedly[17] being carried out by GREENE, FRANKL and ZEMEL.

The question arises as to how far the actual values of surface mobility lie above the experimentally-determined lower limit. This could possibly be estimated from an extensive and systematic study over a large temperature range of samples of widely varying impurity content and surface treatment. At this stage only a few general features will be discussed. Provided that the scattering is not entirely specular, a key parameter determining the amount of mobility reduction at the surface is the ratio of the bulk mean free path λ to the space-charge-layer thickness L. For accumulation layers, λ/L is proportional to $\mu_B|N_D-N_A|^{1/2}$. Thus, the lower the temperature and the higher the impurity content, the greater the reduction in mobility is expected to be. This is substantiated by the experimental curves of Fig. 4, where the mobility is seen to decrease with decreasing temperature in accordance with the increasing λ/L values. The experimental results at the three temperatures, for any given v_s in the region of deep accumulation layers, are roughly in the same proportion as the corresponding theoretical values. This may indicate that the measured values are not too far below the actual surface mobilities. The results of Fig. 5, however, do not appear to support this argument, as the measurements at the two temperatures coincide. Two possible explanations for this behavior present themselves. The first requires the presence of undetected surface states whose effect on the measured results counterbalances that due to the two different values of λ/L. Alternatively, the surface of the sample being considered can become less of a diffuse scatterer as the temperature decreases. These two possibilities can also be applied to account for the data in Fig. 6.

In p-type samples (Fig. 7) all the experimental points lie below the relevant theoretical curves. This is not surprising, however, in view of the characteristics of the trapped-charge release process observed in such samples. In n-type samples, the release times associated with all the states observed were far above the resolution time of the apparatus. Contrary to this behavior, the release process in p-type samples exhibited time-constants right down to the limit of the experimental resolution. It appears therefore that the states present in p-type samples are responsible for the low experimental values measured.

Acknowledgements—The authors wish to thank Mr. S. ALBERT for his help in carrying out the measurements and for his contribution to the processing of the data. Thanks are also due to Mr. M. D. EICHENBAUM for his technical assistance.

REFERENCES

1. KINGSTON R. H. and NEUSTADTER S. F., *J. Appl. Phys.* **26,** 718 (1955).
2. GARRETT C. G. B. and BRATTAIN W. H., *Phys. Rev.* **99,** 376 (1955).
3. SCHRIEFFER J. R., *Phys. Rev.* **97,** 641 (1955).
4. HAM F. S. and MATTIS D., reported at *Amer. Phys. Soc. Meeting, Toronto, June* 22–24, 1955.
5. ZEMEL J. N., *Bull. Amer. Phys. Soc.* ser. 2, **3,** 255 (1958).
6. FLIETNER H., *Ann. Phys., Lpz.* ser. 7, **3,** 396 (1959).
7. FRANKL D. R., *Bull. Amer. Phys. Soc.* ser. 2, **4,** 179 (1959).
8. FRANKL D. R., private communication.
9. BROWN W. L., BRATTAIN W. H., GARRETT C. G. B. and MONTGOMERY H. C., *Semiconductor Surface Physics* (edited by KINGSTON R. H.). University of Pennsylvania Press, Philadelphia (1957).
10. BATH H. M. and CUTLER M., *Bull. Amer. Phys. Soc.* ser. 2, **3,** 138 (1958).
11. PETRITZ R. L., *Phys. Rev.* **110,** 1254 (1958).
12. ZEMEL J. N. and PETRITZ R. L., *Phys. Rev.* **110,** 1263 (1958).
13. ZEMEL J. N., *Phys. Rev.* **112,** 762 (1959).
14. ZEMEL J. N. and PETRITZ R. L., *J. Phys. Chem. Solids* **8,** 102 (1959).
15. MILLEA M. F. and HALL T. C., *Phys. Rev. Letters* **1,** 276 (1958).
16. MANY A. and GERLICH D., *Phys. Rev.* **107,** 404 (1957).
17. GREENE R. F., this Conference, p. 291.

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 193–199. Printed in Great Britain.

FAST SURFACE STATES IN GERMANIUM AT LOW TEMPERATURES

E. HARNIK, Y. GOLDSTEIN, N. B. GROVER and A. MANY

Department of Physics, The Hebrew University, Jerusalem, Israel

VOL. 14 1960

Abstract—The characteristics of fast states on real germanium surfaces have been investigated at low temperatures. Measurements of the time-constant of relaxation processes involving the interaction between the fast states and the majority-carrier band were carried out on *n*-type samples as a function of temperature and surface potential using a pulsed field-effect technique. These were complemented with measurements of trapped charge density as a function of barrier height at various low temperatures. It is shown that the experimental results at such temperatures can be reconciled with room temperature data only by the assumption that both the energy of the recombination center, with respect to the conduction-band edge, and its capture cross-section for electrons are temperature dependent. The recombination level is found to move closer to the conduction band as the temperature is reduced while its capture cross-section for electrons decreases as $\exp(-E_r/kT)$, where the activation energy E_r is about 0·2 eV. There are also indications of surface traps within 0·06 eV from the conduction band which have very small capture cross-sections for electrons. It has been observed that for values of the surface potential corresponding to strong depletion layers the relaxation phenomena are dominated by a process of enhanced generation of minority carriers, the mechanism of which is not understood.

INTRODUCTION

SIMULTANEOUS measurements of trapped charge density and surface recombination velocity as a function of surface potential have proved to be a useful method for the investigation of the parameters of fast states on real germanium surfaces. According to this method the energy parameter of the recombination center, which can be determined very accurately, was used as a starting point for the analysis of the trapped charge vs. surface potential curve. Such measurements have so far been carried out in the vicinity of room temperatures only. There are many apparent advantages in extending this method to low temperatures. The resolution of the various states on the trapped charge vs. surface potential curve can in this way be increased and the exploration of the energy-gap extended to regions nearer the band edges. At low temperatures, however, the filament lifetime becomes too short to be measured, and one has to look for an alternative kinetic process involving time-constants within the resolution of conventional electronic apparatus. The interaction between surface states and the majority-carrier band readily offers such an alternative. Experiments involving such interactions have been carried out by RUPPRECHT[1, 2].

If the equilibrium between a surface trap and the majority-carrier band is suddenly disturbed in such a way as to decrease the occupancy of the trap with respect to majority carriers, the trap must dispose of the excess charge. If the minority-carrier concentration at the surface is negligible compared with the excess charge involved, this can be achieved only through thermal emission into the majority band, provided that the thermal generation of minority carriers is a slower process. If one assumes capture cross-sections of atomic dimensions such as are found at room temperatures the relaxation time of this process in *n*-type samples becomes measurable for levels separated by 15 kT or more from the conduction band. The temperature dependence of the relaxation time-constant would then yield the energy of the level with respect to the conduction-band edge and also its capture cross-section for electrons. Experimental conditions which would restrict the relaxation process to surface trap–majority-carrier band interaction can be obtained in extrinsic samples at low temperatures with accumulation or

depletion layers. If the levels remain at constant energies with respect to the conduction-band edge, they are expected to move progressively beyond the 15-kT limit as the temperature is lowered and thus become accessible to measurement. In this way one expects to be able to scan a considerable region of the gap near the conduction-band edge by measuring the temperature dependence of the relaxation time-constant for various values of the surface potential. These results could then be complemented with measurements of trapped-charge density vs. surface potential at various temperatures.

With these considerations in mind measurements of the temperature dependence of the relaxation time-constant were carried out for various values of surface potential in n-type material. The results obtained show that the situation is more complex than foreseen. Contrary to expectation, longer relaxation time-constants and larger activation energies are observed for surface potentials corresponding to accumulation layers, and these decrease rapidly as the barrier is brought through the depletion region toward inversion layers. It appears that as the energy bands are bent more and more a further mechanism for the release of trapped carriers becomes operative with increasing activity. From trapped charge vs. surface potential curves, it appears that the distribution of fast states at low temperatures is different from that expected on the basis of room temperature data. At low temperatures most of the traps are found to lie near the conduction-band edge and to possess very small capture cross-sections for electrons. To reconcile the experimental results at low temperatures with room temperature data, it is necessary to assume that the recombination center moves toward the conducton-band edge as the temperature is lowered and that its capture cross-section for electrons is temperature dependent with an activation energy of about 0·2 eV. Indications to this effect have been reported previously.[3]

EXPERIMENTAL METHODS

The relaxation process. The kinetic equation for the interaction between surface traps and the majority-carrier band has been developed by RUPPRECHT[1,2] on the basis of the SHOCKLEY–READ[4] and HALL[5] model and the principle of detailed balance. Accordingly the time dependence of the density n_t of majority carriers (e.g. electrons) in surface states at a single energy level is given by

$$dn_t/dt = rn_s(N_t - n_t) - rn_1n_t,$$

where n_s is the surface concentration of free electrons per cm^3, N_t the number of traps per cm^2, r the recombination factor equal to the product of the thermal velocity v_T and the capture cross-section, and n_1 the free electron concentration at the surface for the particular case when the Fermi level for electrons is located at the trap, i.e.

$$n_1 = N_c \exp[(E_c - E_t)/kT].$$

The solution of this equation remains exponential as long as the following two conditions are fulfilled: $n_s \ll n_1$ and n_t is not too small compared to N_t. The time constant τ of the process in this case is simply

$$\tau = (rn_1)^{-1}.$$

Now, for states lying above the Fermi level, $n_s < n_1$, these conditions can be maintained over a considerable portion of the amplitude of the relaxation process if the following procedure is adopted. An external field is applied to the surface, of such polarity as to increase the charge trapped in the states, and is then suddenly removed. The excess trapped charge thus left in the surface will cause a corresponding decrease in n_s, since it repels majority carriers from the surface. The subsequent emission of this excess trapped charge can be measured as a relaxation of the surface conductance to its equilibrium value prior to the application of the external field. This method has the advantage that the equilibrium value to which the surface conductance relaxes is known *a priori*. This becomes particularly useful when the time-constant is very long.

Trapped charge density measurements. To determine the trapped charge density at low temperatures for barrier heights where the surface mobility reduction is important, the change of surface conductance $\Delta\sigma$ with field-induced charge was measured once under non-equilibrium conditions (with negative voltage pulses applied to the field plate) when all the induced charge is assumed to appear as mobile carriers in the space-charge region, and then again under conditions of equilibrium between the space-charge region and the

VOL. 14 1960

surface states (i.e. with d.c. voltage). Fig. 1 shows the results of such measurements and demonstrates the procedure adopted to interpret the results. The abscissa in linear scale is proportional to the field-induced charge density. The upper curve (circles) represents the non-equilibrium pulsed field-effect data, while the lower curve (dots) is from a conventional d.c. field-effect. Evidently the horizontal distance between the two curves at any value of $\Delta\sigma$ represents the change in trapped-charge density as $\Delta\sigma$ is varied from zero to that value, due account being taken of surface mobility reductions. The corresponding value of the surface potential is determined from the upper curve, by making use of the known relation between surface potential and mobile-carrier density in the space-charge layer. This procedure is described in detail elsewhere.[7] It should be pointed out that in the foregoing analysis only those states whose relaxation processes are slow enough to be experimentally observed in the pulsed field-effect measurements are taken into account. The results of such an analysis are shown at the bottom of the figure, on a non-linear scale as a function of the barrier height v_s in units of kT/q.

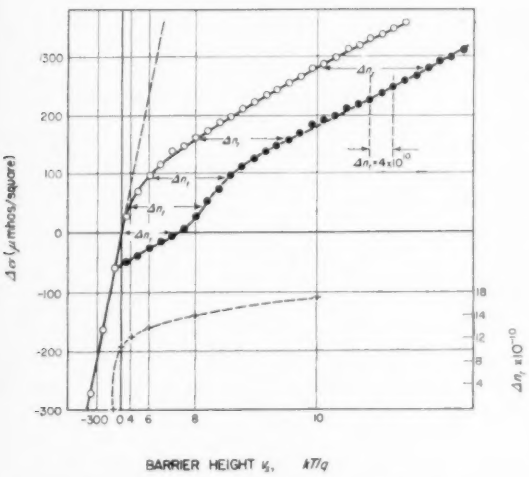


FIG. 1. Typical results of surface conductance $\Delta\sigma$ vs. field-induced charge, used for the evaluation of the trapped-charge density as a function of barrier height (v_s). The upper curve (open circles) represents the results of measurements carried out under non-equilibrium conditions when all the induced charge is supposed to remain mobile. The lower curve (closed circles) is a conventional field-effect curve under conditions of equilibrium between trapped and mobile charges. The horizontal distances defined by arrows are proportional to the change in trapped charge (with respect to $\Delta\sigma=0$), and correspond to the barrier heights indicated. The results of the analysis are also shown (crosses). $T=80°K$.

In some cases an alternative method was used to check the results. A positive voltage of known amplitude was applied to the field plate. The charge thereby trapped was measured directly from the initial value of the surface conductance at the instant this voltage was removed (Fig. 2b). Use was made of the upper curve, as in Fig. 1, to take into account the reduction in surface mobility and to evaluate the barrier height. The results of such a procedure are shown by the dots in Fig. 5(a), together with the results obtained by the method described first (circles). The agreement between the results of the two types of measurements is satisfactory.

Techniques

The germanium samples studied were rectangular filaments about 10 mm long, 3 mm wide and 0·3 mm thick. Following a chemical etch they were mounted on a jig with thin (10μ) Mylar spacers separating the two larger surfaces from the field plate. The jig was inserted into a cryostat in which the temperature could be controlled between 100 and 300°K. The measurements were carried out *in vacuo* except those at 80°K which were made while the jig was immersed directly in liquid air.

The resistance of the filament was measured on the upper half of a double bridge shown schematically in Fig. 2(a). The equilibrium value of the surface potential was varied by capacitively applying a d.c. bias-voltage to the field plate. A positive voltage, in series with the bias-voltage was applied to the field plate, causing $\Delta\sigma$ to increase. This voltage was then suddenly removed, as a consequence of which the surface conductance decreased beyond its equilibrium value and subsequently relaxed back to its value prior to the application of the positive voltage. This decay was displayed on the oscilloscope used as a null indicator in the Wheatstone bridge, as shown schematically in Fig. 2(b). For time-constants in the range from milliseconds to seconds the relaxation of $\Delta\sigma$ was photographed on the oscilloscope. For shorter time-constants, in the range from several milliseconds to a hundred microseconds, the positive

voltage was replaced by a positive square-wave pulse and the time-constant determined by the use of a modified lifetime bridge,[5] as shown schematically in the lower half of the double bridge. Good agreement was obtained in the region where the two methods overlap.

their logical sequence and with reference to a typical n-type sample ($N_D - N_A = 6{\cdot}6 \times 10^{13}$, etched in CP-4). Room temperature data of trapped-charge density and recombination velocity as a function of surface potential were first taken and analysed.[3] The results are shown in Fig. 3(a).

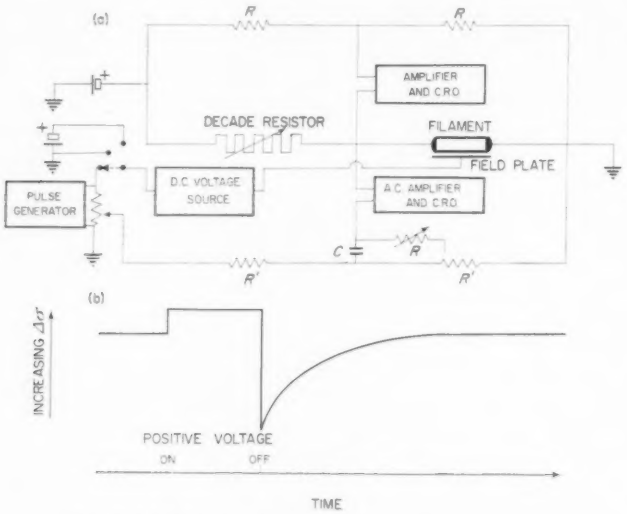


FIG. 2. (a) Schematic outline of the experimental set up. (b) Illustration of the display of the relaxation phenomena.

The measurements of trapped-charge density at low temperatures were carried out as follows. The upper curve in Fig. 1 was obtained by using negative pulses in a method similar to that described elsewhere.[7] To obtain the lower curve shown here, the same voltage as that used for the negative pulses, but in the form of a positive d.c. voltage, was applied to the field plate. The resulting increase in filament resistance was determined by balancing the decade resistor. The d.c. voltage was then removed and the bias-voltage adjusted so as to restore bridge balance. This technique was repeated a number of times. The advantage of this procedure over the conventional d.c. method lies in the equality of the positive-and negative-voltage steps. This increases considerably the accuracy of the subsequent analysis.

EXPERIMENTAL RESULTS

The experimental results will be described in

It can be seen that the recombination center lies at about 0·12 eV ($5kT$ at room temperature) below the conduction-band edge. Its density is $1{\cdot}8 \times 10^{11}$ cm^{-2} and its capture cross-section for electrons is 10^{-13} cm^2. There are no states of appreciable density between the recombination center and mid-gap.

The sample was then cooled and the relaxation phenomena examined for various values of the surface potential and at different temperatures. It was found that, following a rapid initial decay of small amplitude, the relaxation curve was in general exponential down to a few per cent of its maximum amplitude for all values of the surface potential except those corresponding to accumulation layers. A semilogarithmic plot of the time constant vs. $1000/T$ is shown in Fig. 4. The equilibrium value of the barrier height v_s in these measurements remained at all temperatures near $v_s = 0$ (flat bands). Hence only levels lying within $\cong 12\,kT$ from the

VOL. 14 1960

conduction-band edge could be involved in the relaxation process. The fact that the points lie on a straight line for values of the time constant extending over five orders of magnitude proves that only states at a single energy level are involved throughout. The curve implies an activation energy of 0·33 eV. This value cannot, however, be the energy distance between the surface level and the conduction-band edge, since states at such an energy level were not involved in the relaxation process. It follows that part of this activation energy must be attributed to the capture cross-section of the traps for electrons. From room temperature data it follows that the states involved, at least in the higher temperature range of Fig. 4, must be identical with the recombination center. Accordingly these states *are* the recombination centers and hence their capture cross-section for electrons must be temperature dependent with an activation energy of about 0·2 eV. For the purpose of obtaining more information, the trapped-charge density as a function of barrier height v_s was measured at a number of temperatures. The results of two such measurements are shown in Fig. 5. The dots on the lower curve were taken at 233°K and represent the state at the midgap. At this temperature the recombination center does not appear on the curve, as its relaxation is too fast to be observed. On the upper curve, however, which was taken at 97°K, the recombination center *should* appear and it is expected to lie at $qv_s \cong -3\,kT$, as at this barrier height the Fermi level is located at the trap. No state of comparable density is found at this value of v_s. The only relevant state lies at $qv_s \cong 5\,kT$ (0·05 eV below the edge of the conduction band). It appears, therefore, that the recombination level moved closer to the conduction

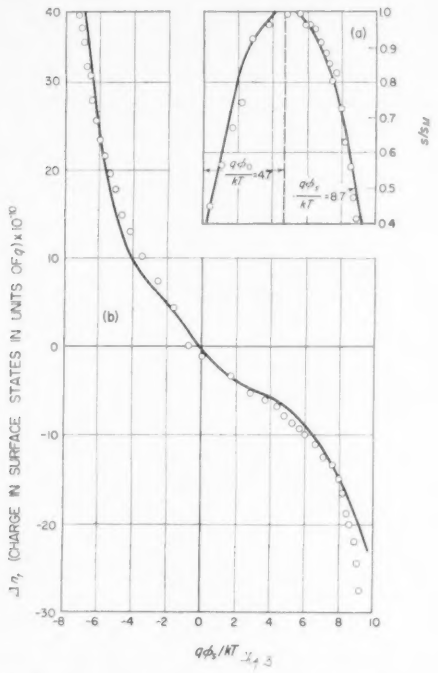


FIG. 3. Surface recombination velocity (a) and trapped charge density (b) vs. $q\phi_s/kT$ for an n-type sample ($N_D-N_A = 6{\cdot}6\times10^{13}$, etched in CP-4), at room temperature.

$$(E_t-E_i)/kT = 8{\cdot}7;\ N_t = 1{\cdot}8\times10^{11}$$
$$(E_t'-E_i)/kT = -0{\cdot}3;\ N_t' = 1{\cdot}2\times10^{11}$$
$$(E_t''-E_i)/kT = 6{\cdot}6;\ N_t'' = 8\times10^{11}$$
$$T = 293°\text{K}$$

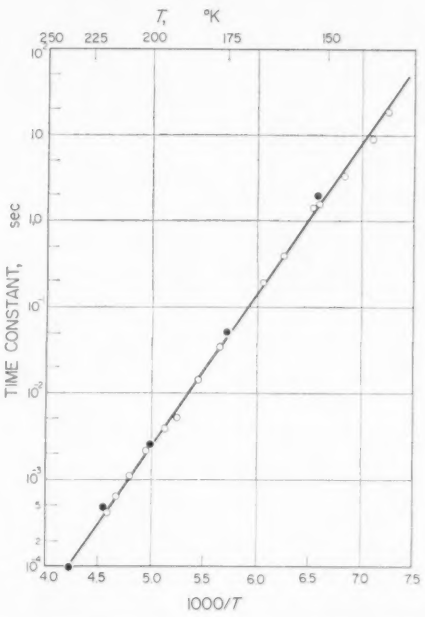


FIG. 4. Semilogarithmic plot of the relaxation time-constant vs. $1000/T$ for an n-type sample ($N_D-N_A = 6{\cdot}6\times10^{13}$, etched in CP-4), in the vicinity of $v_s = 0$. Activation energy = 0·33 eV. ● Plate voltage = +25 V ($v_s \simeq +1$), ○ Plate voltage = 0 ($v_s \simeq -1$).

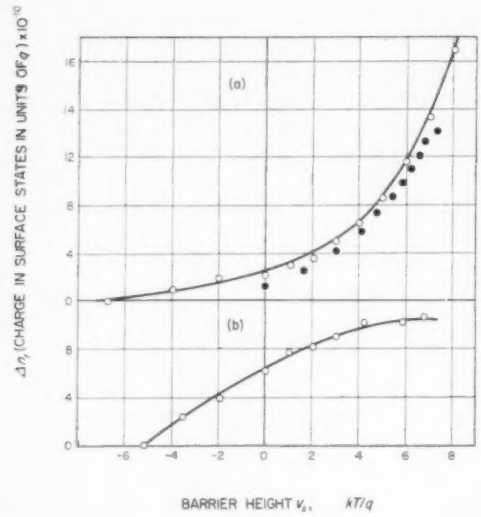


FIG. 5. Trapped-charge density vs. barrier height (v_s) for an n-type sample ($N_D-N_A = 6{\cdot}6\times10^{13}$, etched in CP-4), at 97°K (a) and at 233°K (b). The closed circles in (a) represent results obtained by an alternative procedure as described in text.

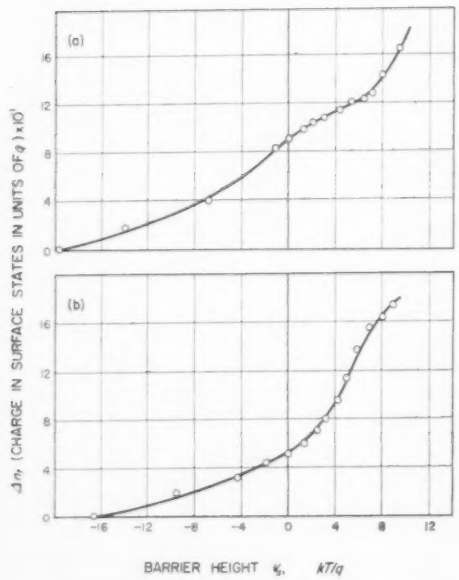


FIG. 6. Trapped charge density vs. barrier height (v_s) for an n-type sample ($N_D-N_A = 4{\cdot}4\times10^{13}$, subjected to warm H_2O_2 treatment), at 80°K (a) and at 152°K (b).

band as the temperature was lowered. This conclusion is in accordance with previously reported results.[3] Fig. 6 shows curves of trapped charge density vs. barrier height for another sample ($N_D-N_A = 4{\cdot}4\times10^{13}$, treated with warm H_2O_2). In this case the room temperature measurements placed the recombination center at 0·2 eV from the edge of the conduction band. Again there is no comparable density of states at barrier heights where the recombination center is expected to lie ($qv_s \cong -4\,kT$ at 152°K and $qv_s \cong -17\,kT$ at 80°K).

The relaxation-time constant was found to decrease as the surface potential was decreased.

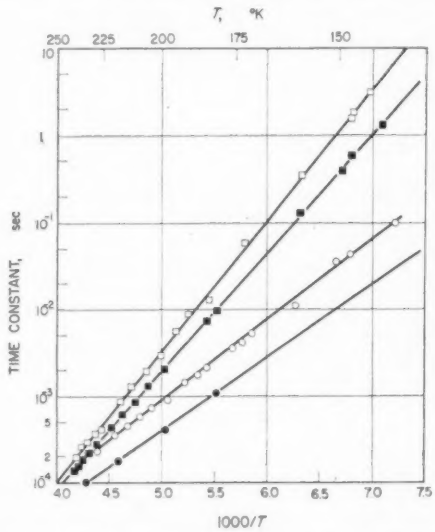


FIG. 7. Semilogarithmic plots of the relaxation-time constant vs. $1000/T$ in an n-type sample ($N_D-N_A = 6{\cdot}6\times10^{13}$, etched in CP-4), for various values of the d.c. bias, i.e. for different values of the surface potential.

	Plate voltage (V)	Activation energy (eV)
□	−100	0·28
■	−200	0·24
○	−400	0·15
●	−1000	0·13

Semilogarithmic plots of the relaxation-time constant vs. $1000/T$ are shown in Fig. 7 for values of the equilibrium surface potential corresponding to strong depletion layers. It can be seen that the activation energy, as determined from the slope of the curves (corrected for the temperature dependence of N_c, and V_T), decreases progressively

VOL. 14 1960

as the surface potential decreases. These activation energies cannot be associated with simple thermal emission from surface states, since no levels corresponding to these activation energies were found either at room temperature or below. They appear rather to be connected with the emptying of deeper levels through an enhanced generation of minority carriers in the surface region, the exact mechanism of which is not yet understood. This possibility is further supported by the characteristics of extraction phenomena which make their appearance as the surface conductance minimum is approached during measurements using the lifetime bridge. The generation-rate of minority carriers, as derived from the time-constant of the exclusion transient, follows roughly the trend shown by the relaxation-time constant as the surface potential is varied.

CONCLUSIONS

From the experimental results presented it appears that the electrical behavior of real germanium surfaces is more complex than could have been predicted from room temperature data. A self-consistent model of the surface can be formed only if it is assumed that both the energy and the capture cross-section for electrons of the recombination center are temperature dependent. As the temperature is lowered the recombination center moves closer to the conduction band, while its capture cross-section for electrons decreases exponentially with $1/T$. Accordingly no new states are involved in the experimental results described. The relaxation processes shown in Fig. 4 involve at all temperatures the interaction between the recombination center and the conduction band. The experimentally determined activation energy is no longer a direct measure of the energy level of the states, but includes a contribution due to the temperature dependence of the capture cross-section. In order to account for these results, the relaxation time-constant must be written in the form

$$(\tau)^{-1}=r_0 \exp(-E_r/kT)\times N_c \exp[-(E_0+\alpha T)/kT]$$

where, for the sample discussed, $E_r \cong 0{\cdot}20$ eV, $E_0 \cong 0{\cdot}03$ eV and $\alpha \cong 3\times 10^{-4}$ eV/deg.

The temperature dependence of the capture cross-section is not altogether surprising. The recombination center is an acceptor-type level, and such levels are often found to have temperature-dependent cross-sections for electrons with activation energies of this order of magnitude.[8] This temperature dependence implies that the recombination center is negatively charged when empty. The temperature dependence of the energy of the recombination center, however, is more difficult to understand.

Data of trapped-charge density at room temperature show that there are some unidentified surface levels near the conduction-band edge. The appearance of such states on the trapped-charge density curve at 80°K (Fig. 6a) indicates that their capture cross-section for electrons must be very small.

The possibility that some or all of the relaxation processes observed involve bulk traps rather than surface levels has been pointed out by RUPPRECHT[2-3]. Such bulk effects are negligible under the experimental conditions employed here, as the surface density of the bulk traps involved could not have exceeded 10^9 cm^{-2}. It is quite probable, however, that the rapid initial transients of low amplitude observed on the relaxation curves are due to just such bulk effects. The temperature dependence of the time constants of these transients indicates a small activation energy of about 0·1 eV. This is fairly consistent with that expected from bulk traps, in view of the position of the Fermi level in the bulk.

Acknowledgements—The authors wish to thank Mr. S. ALBERT for his help in carrying out the measurements and for his contribution to the processing of the data. Thanks are also due to Mr. M. D. EICHENBAUM for his technical assistance.

REFERENCES

1. RUPPRECHT G., *Phys. Rev.* **111,** 75 (1958).
2. RUPPRECHT G., *Bull. Amer. Phys. Soc.* **3,** 377 (1958).
3. MANY A. and GERLICH D., *Phys. Rev.* **107,** 404 (1957).
4. SHOCKLEY W. and READ W. T., *Phys. Rev.* **87,** 835 (1952).
5. HALL R. N., *Phys. Rev.* **83,** 228 (1951).
6. MANY A., *Proc. Phys. Soc., Lond.* **B67,** 9 (1954).
7. MANY A., GROVER N. B., GOLDSTEIN Y. and HARNIK E., this Conference, preceding paper.
8. BAUM R. M. and BATTEY J. F., *Phys. Rev.* **98,** 923 (1955); BATTEY J. F. and BAUM R. M., *Phys. Rev.* **100,** 1634 (1955).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 200–207. Printed in Great Britain.

SURFACE RELAXATION EFFECTS IN GERMANIUM AT REDUCED TEMPERATURES

D. H. LINDLEY and P. C. BANBURY

Physics Department, University of Reading, Reading, England

Abstract—Measurements of small signal field effect mobility on germanium as a function of frequency of the applied a.c. field and under-pulsed fields have been extended in temperature down to 100°K. The measurements have been made for various values of a superimposed d.c. field giving a range of barrier heights as deduced from the d.c. conductance. Relaxation processes have been observed with characteristic times in the range $1–10^{-6}$ sec, depending on the temperature and the barrier height. In inversion layers the relaxation can be represented by a single time constant, whilst in other types of barrier the relaxations are not exponential. The effects are sensitive to illumination. The results are discussed in terms of surface trapping and recombination, and a comparison is made with predictions from GARRETT's analysis.

VOL
14
196

1. INTRODUCTION

FAST surface state properties associated with the screening of charge induced in conducting states in the semiconductor and surface recombination velocity, have been investigated in detail by a number of workers. Implicit in the models used is the process of trapping of carriers in a surface state from either band at the surface. This process has not been directly studied in most experiments to date. RUPPRECHT[1] has observed decays in the conductance of a pulsed channel as a function of temperature and MORRISON[2] the decays of conductance on application of a large electric field at liquid nitrogen temperatures. MONTGOMERY[3] observed at room temperature a relaxation in the region of 5×10^7 c/s which he attributed to trapping. Following MONTGOMERY's work, GARRETT[4] analysed the behaviour of the conductance of a germanium filament as function of a small applied electric field of variable frequency. He treated the model of a single surface state obeying SHOCKLEY–READ[5] statistics and shewed that two dispersions in conductance were to be expected. One was connected with the time for generation and recombination of minority carriers in the filament arising from extraction and injection across the barrier bulk interface. The other was essentially the time for carriers to be trapped by the state from the valence or conduction band. An extension of the analysis has been made by BERZ[6]. MONTGOMERY[3] and YUNOVICH[7] have made qualitative measurements on the lifetime dispersion. The present authors[8] have made more detailed measurements and have found substantial agreement at room temperatures with GARRETT's predictions, and with measurements made by other workers using different techniques. Since the room temperature trapping time is too short to detect by our experimental procedure, we have made measurements of field effect dispersion and relaxation effects at lower temperatures, and have compared the predictions of the model with regard to variations both with temperature and surface potential.

2. THE MODEL

GARRETT's analysis indicates two relaxations which are interdependent in the sense that they are both a function of the surface state parameters and the surface potential. This need not necessarily be the case since a centre which screens induced charge is not *a priori* a significant recombination level. In this discussion we will assume they are identical but the general conclusions will be largely unaffected by this assumption. We shall be mainly interested in the predictions as a function of temperature.

Under certain conditions, fulfilled in our experiments, the dispersion associated with the generation

and recombination of minority carriers is the filament lifetime τ_f as defined by SHOCKLEY[9] where

$$1/\tau_f = 1/\tau_b + s/T + s^*/T$$

Here τ_b is the bulk lifetime, T the thickness of the slab and s and s^* are the surface recombination velocities at the front and back faces. s is given by

$$\frac{(n_0+p_0)C_pC_n}{C_p(p_s+p_1)+C_n(n_s+n_1)} \qquad (1)$$

where n_0, p_0 are the bulk carrier concentrations, n_s, p_s are the surface carrier concentrations, p_1, n_1 are the carrier concentrations if the Fermi level were at the trap energy and C_p, C_n are the hole–electron capture cross sections. The temperature dependence of this dispersion time constant is simply that of the lifetime of an injected electron hole pair, which decreases with decreasing temperature. For a thin filament of reasonably long bulk lifetime this is dependent on the sample geometry and the magnitudes and temperature dependence of s and s^*.

The trapping time τ_t or surface state relaxation time is independent of sample dimensions and back surface and is a true relaxation time in the sense that it is not an averaged behaviour over the whole filament. It is given by

$$\tau_t = \frac{1}{\Omega}\,\frac{\partial/\partial u_s(\Gamma_p-\Gamma_n)}{\partial/\partial u_s(\Gamma_p-\Gamma_n)-N_t f_0(1-f_0)} \qquad (2)$$

where

$$\Omega = [C_p(p_s+p_1)+(n_s+n)C_n]/N_t$$

N_t is the trap density, f_0 the Fermi occupancy factor at the trap energy and u_s is the dimensionless variable related to surface potential defined by KINGSTON and NEUSTADTER[10].

This clearly has the opposite temperature dependence of τ_f; i.e. it becomes longer as the temperature decreases. The dependence both as a function of temperature and of surface potential u_s is to an approximation that of s.

Substituting experimental values from room temperature measurements the trapping time is found to be of order 10^{-7}–10^{-8} sec. As the temperature is lowered τ_f and τ_t should approach each other and at some temperature cross. Below this temperature in regions of barrier height where contributions to conductance by minority carriers are negligible (i.e. accumulation and exhaustion layers), the only relaxation observable at small disturbances should be τ_t. In intrinsic material this region covers most of the centre of the gap at low temperatures. Since τ_t is longer than the lifetime the carrier concentrations in the bulk and barrier are closely those of equilibrium during the relaxation. This replaces GARRETT's assumptions about diffusion currents from the bulk to the barrier. In inversion layers at extreme barriers we expect more complicated behaviour at room temperature since lifetime relaxations will still be important in addition to possible surface state trapping.

In the regions of surface potential where the surface states under discussion are responsible for all screening of induced charge the field effect mobility, μ_{fe}, defined as

$$\Delta G_s/\Delta\sigma$$

(ΔG_s is the surface conductance change per square for applied charge per unit area $\Delta\sigma$) at times short in comparison with τ_t will be closely that of the bulk majority carrier. In inversion regions following GARRETT's room temperature discussion for times short in comparison to τ_f and long compared to τ_t we expect μ_{fe} to exceed the bulk mobility

Since trapping is a process determined in magnitude by densities of unoccupied states and carrier concentrations in the conduction and valence bands we may forecast that τ_t will be sharply influenced by creation of hole electron pairs at the surface, e.g. by illumination; τ_f should be much less influenced.

The above discussion rests on the assumption that the induced charge disturbances lie in the small signal range.

3. EXPERIMENTAL RESULTS

(a) *Experimentation*

Field effect relaxation measurements have been made on CP-4 etched germanium filaments of dimensions of the order of $1{\cdot}5\times0{\cdot}3\times0{\cdot}02$ cm both as a function of temperature and surface potential u_s. The slabs were mounted in a jig with a large face against a lapped metal electrode with dielectric spacer of melinex or mica, and the jig

suspended in a glass cryostat. The conductivity changes were measured by a probe method or using the current carrying contacts as voltage probes. Similar results were obtained in all configurations. Since slow state relaxations are of order of minutes or hours at low temperatures, u_s can be held at different fixed values by application of a variable large d.c. field. Thus in regions where a conductance minimum can be obtained and hence u_s evaluated, the surface relaxations may be investigated as a function of u_s by superimposing on the d.c. field a small a.c. field of variable frequency, a pulse or a square wave. The circuit for separation of displacement current and conductivity modulation signals has been described elsewhere.[8] Decays and dispersions in the range 1–10^{-6} sec could be measured. Most of the detailed measurements have been made in the range 210°–230°K for n-type material and 230°–170°K for p-type since below these regions a minimum could not be obtained and above these regions the slow states are too fast.

Facilities were also arranged for shining light from a tungsten bulb on the field effect face by use of a transparent field electrode. As no effort was made to achieve uniform illumination these results must be regarded as roughly qualitative only.

(b) *Results*

Figure 1 shows typical results of pulsed field and a.c. field effect measurements. In exhaustion-accumulation layers the relaxations cannot be represented by a single relaxation time τ. However, in most cases they can be approximated to one. We shall, for convenience, assign time constants to our results though strictly they are not meaningful.

At small disturbances the transient decays are symmetrical with sign of the induced charge. At higher fields the decays become asymmetrical. Fig. 2 shows a decay as a function of size of charge induced in an n-type specimen. The decay corresponding to increase of minority carrier in the barrier becomes markedly longer and non-exponential with increasing amplitude of induced charge. This asymmetry tends to become more pronounced as the temperature decreases and is reversed for p-type material, i.e. the longer decay still corresponds to increase of minority carriers. The decay time corresponding to decrease of minority carrier concentration in the barrier is nearly independent of amplitude at low levels of induced charge. The a.c. dispersion tends to average the two decays at large disturbances. At low fields the two measurements agree. Since most of the measurements of τ were taken at induced charge of 10^9 electrons/cm^2 or less corresponding

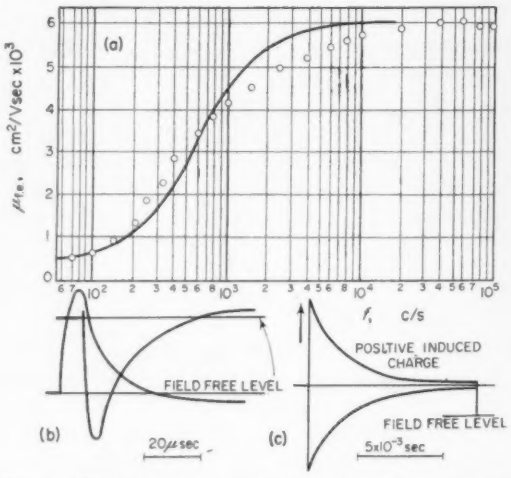


FIG. 1. Typical small signal dispersion and decays on n-type specimens. (a) was measured at 220°K. The theoretical curve is drawn for a single time constant dispersion of characteristic frequency $6{\cdot}10^2$ c/s. (b) and (c) are traced from oscilloscope photographs and show decays for both signs of induced charge. (b) is for an inversion layer at $T = 247$°K. (c) is at $T = 198$°K and $u_s \simeq 0$.

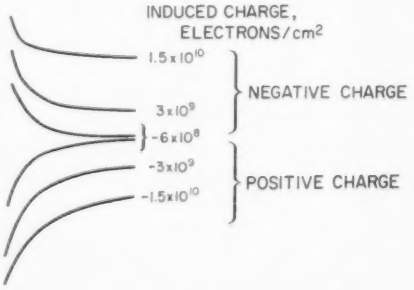


FIG. 2. Decays for an n-type specimen as a function of the amplitude of induced charge. $T = 220$°K. The decays for 6×10^8 electrons/cm^2 correspond to an initial change in surface potential of about $0{\cdot}75\ kT$. The top three curves represent a decrease of hole concentration in the barrier, the lower three an increase in concentration.

to the range of approximate agreement, no specific mention will be made as to the method of evaluating τ.

In inversion layers the decays are good approximations to exponentials (Fig. 1) and tend to be symmetrical over a considerable induced charge range .They correspond to a change in sign of the field effect and hence must be associated with a lifetime effect.

Figure 3 is a plot of relaxation time, τ, versus temperature for n and p-type specimens to below 200°K. Three of the plots are for inversion layers.

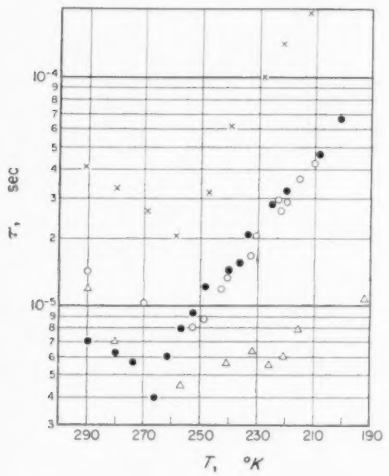


FIG. 3. Relaxation time τ against temperature. Points (a) ×, (b) ○, (c) ● are for n-type specimens (d) △ for p-type. The surface of (a) is in an exhaustion–accumulation layer condition. (b) and (c) are for inversion layers and represent the same field effect surface but in (c) the back surface has been sandblasted. (d) is also for inversion layers.

Curve (c) is for the same surface as (b) but with the back surface sand-blasted instead of etched. The other plot (a) is for a non-inversion layer. The numerical magnitudes of u_s, which change continuously with change of temperature, are undetermined. The main point of interest is the sharp break in temperature variation of τ at around 260°K.

Figure 4 is a graph of measured relaxation time, τ, versus u_s for three different n-type specimens at about 220°K. The variations found are always quantitatively similar. Differences usually lie in the extent of the flat top of the curve. Sometimes there is a sharp peak and decrease of τ into accumulation layers—sometimes no drop at all. Also drawn in are two theoretical plots, for a discrete surface state at $E_t - E_i = -4\,kT$ and $E_t = E_i$.

In Fig. 5 is shown a typical plot of τ versus u_s for a p-type specimen at 178°K. The inversion layer time constants are different in order of magnitude from those in accumulation and exhaustion regions. Also shown are results for an n-type specimen at 198°K with a theoretical curve $E_t - E_i = -2\,kT$. The values of u_s are less reliable

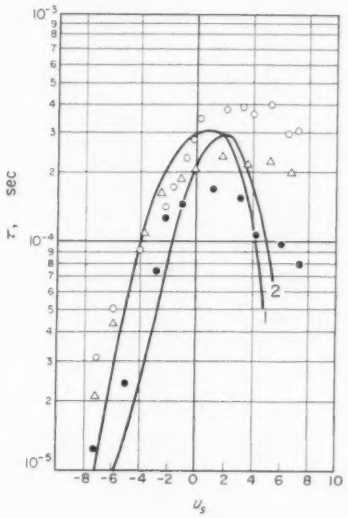


FIG. 4. Relaxation time τ against surface potential u_s for three n-type specimens at 220°K. The theoretical curves are drawn for (1) $E_t - E_i = -4kT$, $N_t = 2\times10^{10}/\text{cm}^2$, $C_p = C_n = 0{\cdot}910^{-8}$, $N_t$cm/sec. (2) $E_t = E_i$, $N_t = 2\times10^{10}/\text{cm}^2$, $C_p/N_t = C_n/N_t = 1{\cdot}7\times10^{-9}$ cm^3/sec.

as the minimum of conductance had disappeared from the range of measurement. The relative values of u_s in the two cases are not directly comparable since the temperatures are different. Qualitatively the results for the two materials are similar but the relaxation times for comparable temperatures are in general longer for n-type than for p.

In addition to the above curves in Fig. 5 are two plots of τ under illumination. Under extremely weak illumination τ decreases with increase in intensity, but the variation with us tends to remain

the same as for the unilluminated results. At higher illuminations τ tends to saturate with intensity and become independent of u_s. The transient decays for either sign of induced charge are equally affected by the illumination.

Figure 6 shows typical data of μ_{fe} versus u_s for an n-type specimen for times shorter and longer than the relaxation time τ. The bulk electron mobility of $6{\cdot}2 \times 10^3$ cm^2/V sec is drawn in. It can be seen that over a substantial region between $u_s = 6$ and $u_s = -6$, $\mu_{fe} \cong \mu_n$. Included in the same figure is μ_{fe} for illumination on the same surface at 220°K.

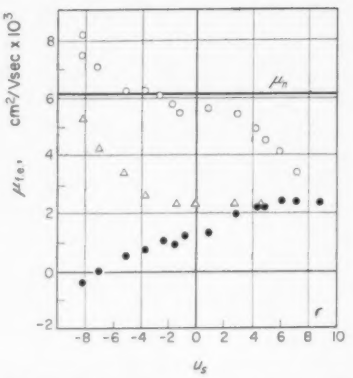


FIG. 6. Field effect mobility $\mu_{f.e.}$ against surface potential u_s for an n-type filament at 220°K. Points ○ are for times short in comparison to the trapping time τ. Points ● are the equilibrium values, i.e. at times long compared to τ. Points △ are for short times but under strong illumination. The bulk electron mobility of $6{\cdot}2 \times 10^3$ cm^2/V sec is indicated.

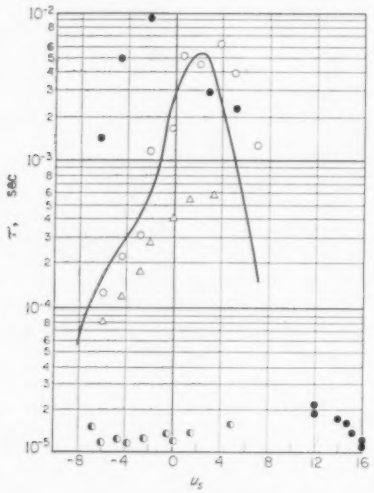


FIG. 5. τ vs. u_s for an n-type specimen at $T = 198$°K, for a p-type specimen, $T = 178$°K, and for an n-type under strong and weak illumination. The theoretical curve is drawn for $E_t - E_i = -2kT$. $N_t = 5 \times 10^{10}$/cm^2, $C_p/N_t = C_n/N_t = 1{\cdot}1 \times 10^{-9}$ cm^3/sec.

○ n-type, 198°K.
● p-type, 178°K.
△ n-type, 198°K, weak illumination.
◐ n-type, 200°K, strong illumination.

Fig 7. shows a μ_{fe} curve for a p-type specimen at 178°K. These figures show that for inversion layers on both n and p samples, μ_{fe} at times shorter than τ, becomes considerably greater than the bulk mobility of the majority carrier.

Finally we have looked at highly developed n and p accumulation layers for trapping in states close to the band edges down to 100°K and failed to find any decays up to times of order 10^{-6} sec.

Also, measurements by CAJKOVSKI in this laboratory[11] of the change in trapped charge with u_s down to temperatures of 170°K have failed to show signs of any structure to be associated with a "discrete state". We shall make use of these two additional pieces of evidence in our following discussion.

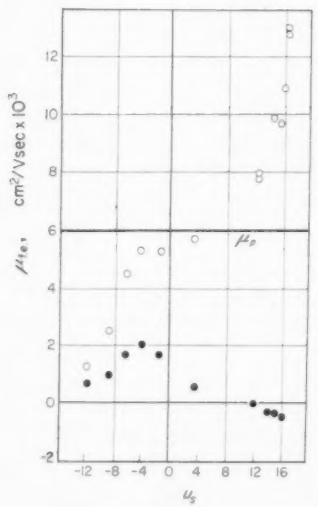


FIG. 7. $\mu_{f.e.}$ vs. u_s for a p-type specimen at $T = 178$°K. Points ○ are for times shorter than the trapping time τ. Points ● are the equilibrium values. The bulk hole mobility $\mu_p = 5{\cdot}95 \times 10^3$ cm^2/V sec is drawn in.

VOL. 14 1960

4. CONCLUSIONS

As is clear from the comparison of Sections 2 and 3(b) there are fairly well-defined areas of agreement and disagreement between the expected and observed behaviour.

The decays and dispersions will be treated as if having a single time constant in the main discussion and a comment made later on this approximation. We shall also assume from empirical evidence that data from transient phenomena may be treated on an equal footing with steady state measurements.

The variation of τ with temperature in an exhaustion region is as predicted from the model. The transition region is at a higher temperature than would be expected from room temperature cross sections and recombination energy levels. However, we are not bound to associate the low temperature decays with a recombination centre. The decays at low disturbances (of the order of 1–2 kT and less) are as expected, symmetrical with the sign of the induced charge. At higher induced charge the decay corresponding to the increase of minority carriers in the barrier becomes very much longer than that for the decrease. Following arguments given by BANBURY *et al.*[12] and MORRISON[2] we can attribute this to the small generation rate in the filament at low temperatures.

We may regard the generation and the recombination processes at the surface as represented by

$$g - r = \text{net generation rate}$$

where to an approximation, when the Fermi level lies near the centre of the gap, g is a constant generation rate proportional to n_i^2 and r is the recombination rate. When the disturbance becomes large in the limit the largest net generation rate afforded to the system is g while the recombination rate can increase indefinitely. At low temperatures g is small and the largest possible diffusion currents into the barrier from the bulk are negligible since the bulk minority carrier concentration is small. The reverse processes of recombination and diffusion into the bulk are however not restricted in this way. The above simple arguments would suggest that at large disturbances an asymmetry would be expected in the transient decays resulting from inducing positive and negative charge in the semiconductor at low temperatures in regions of barrier height where minority carriers are important either in conductance or in trapping in surface states. Since the swing in u_s under large applied field in Fig. 2 includes the regions of u_s where a surface state at the Fermi level would trap holes the generation rate in this case is undoubtedly becoming the rate limiting step in the decay. At very large disturbances the decay should tend to be linear in time, i.e. proportional to the fixed generation rate, as observed by MORRISON. In the other direction of induced charge the rate of decay is controlled by the capture rate for holes or emission rate for electrons in n-type material. This simplified model is felt to give some indication of the situation. The asymmetry could be expected to be more marked as the temperature is reduced and to reverse with n- and p-type material as is observed.

The τ versus u_s data shown in Figs. 4 and 5 follow qualitatively the dependence of the theoretical curves. A discrete level however is not sufficient to account for the results since the range of u_s over which it may screen is not large enough to account for the variation of trapped charge with u_s[11] and hence the full curve of τ_t versus u_s cannot be invoked. Nor does the usual room-temperature surface recombination level seem appropriate since the amplitude of the dispersion is small in that region for n-type specimens (Fig. 6.). The theoretical curves drawn are not therefore intended as a fit but as an illustration of general trends. They suggest that the carrier concentrations in the bands n_s, p_s dominate the trapping time in the expression for τ_t. If there are various sets of discrete trapping levels throughout the gap with differing capture cross sections then the magnitudes of the latter for the relevant carrier are sufficiently similar for n_s, p_s to dominate.

The rough values of capture cross section implied by the results are of interest. The trapping times for n- and p-type materials are often an order of magnitude different at a given temperature. As measurements on p-type material have been less detailed, we restrict our discussion to n-type specimens pending further evidence on this difference.

If we were to assume that there is a trap at the centre of the gap with equal capture cross sections for electrons and holes, then approximating the term multiplying Ω^{-1} in equation (2) to unity,

C_n/N_t at 220° would be 10^{-8} cm^3/sec, a value of reasonable magnitude in comparison with room temperature recombination data. The values of trapping times at other temperatures would be simply governed by the change in n_i: hence at 198°K, $\tau_t = 5{\cdot}10^{-3}$ sec—close to the value observed. At 260°K, $\tau_t = 25{\cdot}10^{-6}$ sec—a reasonable value for the filament lifetime also and therefore consistent with a transition from τ_f to τ_t at this temperature.

From the above it could be argued that traps near the centre of the gap are suggested by the results. It seems difficult to accept a patch effect large enough to provide an alternative explanation although some surface inhomogeneity is indicated by the consistent non-exponential nature of the decays even at small disturbances. To obtain any consistency with surface recombination measurements at room temperature it would be necessary to interpret the latter in terms of the possible level near the centre of the gap rather than the one usually selected (i.e. about 0·12 eV above the gap centre). In addition we fail to find trapping times of the order of 10^{-5}–10^{-6} sec at $T \cong 200$°K implied by the accepted energy.

The trapping time at room temperature for our selected level would be $\sim 2 \times 10^{-6}$ sec. However, we have ignored the correction factor in equation (2) which could be as large as 10 and any shift from the centre of the gap would tend to reduce the time.

Although our experiments are suggestive in the above manner, we think it of little use to try and evaluate numerically the surface state parameters from the results.

The behaviour of the relaxation times in inversion layers is not well understood. Since in this region of u_s at low temperatures the band edge states are operative in screening, we will assume that if n- and p-type germanium have similar surface state structures near the band edges then the trapping times of the states are very much less than 10^{-6} sec in the range of temperature investigated. The relaxations seen are then postulated as some type of "lifetime" relaxation. This is confirmed by the observation that they are always good exponentials and that the field effect mobility is greater than the bulk mobility (Figs. 6 and 7) as might be expected from GARRETT's analysis, and observations at room temperatures.[3, 4, 8] Additional evidence is the large separation of the time constants in an inversion layer and in the exhaustion-accumulation regions for the p-type specimen in Fig. 5. However, we are faced with the apparent contradiction that the decays are more strongly symmetrical in this region than in the centre of the gap (where the surface generation rate is expected to be higher) and that the temperature dependence below 260°K is wrong for a lifetime process (Fig. 3). Moreover, when the back surface of an n-type specimen was sandblasted, the decay at room temperature was reduced (Fig. 3) but below the transition point the times were as before, indicating that the back surface did not take part in the process. Since the diffusion length was greater than the thickness of the specimen from measured τ_b as a function of temperature this too cannot be understood. It would seem from the occurrence of a transition in the temperature dependence that the process involved is connected with the front surface trapping phenomenon, but as yet we are unable to offer a detailed explanation.

The illumination effects are qualitatively understood as a reduction in trapping time by the increased carrier concentrations at the surface. The trapping time decreases with increase in intensity until it becomes comparable with and shorter than the filament lifetime. When this occurs the relaxation observed is the filament lifetime as at room temperatures. The results in Fig. 5 bear this out. This interpretation is additionally strengthened by the large decrease in μ_{fe} at times shorter than τ as shown in Fig. 6. This plot is characteristic of a lifetime dispersion and represents the minority carrier conductance contribution in the barrier.

Acknowledgements—Thanks are due to Professor R. W. DITCHBURN for his encouragement and for the provision of research facilities, to D. CAJKOVSKI for valuable discussions and for access to unpublished work, and to the Admiralty for support and permission to publish this paper.

REFERENCES

1. RUPPRECHT G., *Phys. Rev.* **111,** 75 (1952).
2. MORRISON S. R., *Phys. Rev.* **114,** 437 (1959).
3. MONTGOMERY H. C., *Phys. Rev.* **106,** 441 (1957).
4. GARRETT C. G. B., *Phys. Rev.* **107,** 478 1957.

5. SHOCKLEY W. and READ W. T., *Phys. Rev.* **87,** 835 (1952).
6. BERZ F., *J. Electr. Contr.* **6,** 97 (1959).
7. YUNOVICH A. E., *J. Tech. Phys., Moscow* **28,** 689 (1958).
8. LINDLEY D. H. and BANBURY P. C., *Proc. Phys. Soc., Lond.* **74,** 385 (1959).
9. SHOCKLEY W., *Electrons and Holes in Semiconductors,* p. 315. Van Nostrand, New York (1950).
10. KINGSTON R. H. and NEUSTADTER S. F., *J. Appl. Phys.* **26,** 718 (1955).
11. CAJKOVSKI D., to be published.
12. BANBURY P. C., CAJKOVSKI D. and NIXON J. D., *Bull. Amer. Phys. Soc.* ser 2, **3,** 376 (1958).

VOL.
14
1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 208–213. Printed in Great Britain.

CROSS SECTIONS OF MIDGAP SURFACE STATES IN SILICON BY PULSED FIELD EFFECT EXPERIMENT

G. RUPPRECHT

Research Division, Raytheon Company

Abstract—The cross sections and energy levels of surface states in silicon around midgap have been investigated by a pulsed field effect technique. Two energy levels have been found, one acceptor-like (A), the other one donor-like (D).

The results are: $(E_C-E_A) = 0{\cdot}52$ eV, $\sigma_n^A = 3\times10^{-16}$ cm^2
$(E_A-E_V) = 0{\cdot}72$ eV, $\sigma_p^A = 4\times10^{-12}$ cm^2
$(E_D-E_V) = 0{\cdot}54$ eV, $\sigma_p^D = 2\times10^{-15}$ cm^2
$(E_C-E_D) = 0{\cdot}63$ eV, $\sigma_n^D = 1\times10^{-13}$ cm^2

VOL
14
196

INTRODUCTION

THE existence of dominant surface states in the middle of the energy gap in silicon has been suspected for a long time. The observed stability of silicon surfaces gave support to this assumption. Field effect experiments lead to the conclusion that the density of these states must be appreciable. The channel effect experiment, the most fruitful tool in the investigation of silicon surface states so far, can hardly be applied to these midgap levels because the channel conductance would be too small.[1] At the present time, apparently very little is known about states in the middle of the gap in silicon, such as their charge character, their accurate position and their density. Using pulsed field effect technique, more detailed information can be obtained. Experiments have been carried out on both n-type and p-type silicon samples with different carrier concentrations. Two discrete levels close to the middle of the gap were observed and it was shown that these states exist predominantly at the surface.

PULSED FIELD EFFECT

With a conventional a.c. field effect experiment it is possible to decide whether the surface of a semiconductor is n-type or p-type. Since the position of the Fermi level in the bulk of a sample can be calculated from Hall effect and conductivity data, the bending of the bands towards the surface is essentially known. If a surface charge Q_0 is assumed, a charge Q_S in the space charge layer insures charge neutrality. (The Q's are measured in C/cm^2.) The space charge layer can be considered as a capacitor, C_S, which is voltage dependent. In series with C_S is the capacity, C_P, located between the field plate and the semiconductor surface. The thickness of the dielectric spacer is L and the dielectric constant, ϵ_P. The total capacity per unit area is:

$$C = C_P\left(1+\frac{C_P}{C_S}\right)^{-1} \text{ with } C_P = \frac{\epsilon_P\epsilon_0}{L}.$$

If the voltage V is applied to the field plate, the total charge becomes

$$Q_T(t,V) = C_P\cdot V\left(1+\frac{C_P}{C_S(V)}\right)^{-1}$$
$$= Q_S(t,V)+Q_0(t,V). \qquad (1)$$

Q_S is the total charge in the space charge layer: Q_0 the charge in surface states. C_S can be calculated in the Schottky approximation if no inversion layer is involved. In this case the error introduced in equation (1) by the Schottky approximation is small. For a small surface potential, when C_S is large and the approximation starts to fail, the error in Q_T still stays small, since $C_P/C_S \ll 1$. When C_S becomes smaller with increasing bending of the energy bands, the approximation gets increasingly better.

If one assumes that the concentration N of ionized impurities stays constant throughout the bulk,

the surface conductance is proportional to Q_S and it is, therefore, only necessary to consider the transient behaviour of $Q_S(t,V)$. From equation (1) one finds:

$$\frac{dQ_S}{dt}+\frac{dQ_0}{dt} = C_P V\cdot\frac{d}{dt}\left[1+\frac{C_P}{C_S}\right]^{-1}$$
$$= V\left(\frac{C_P}{C_S}\right)^2\left[1+\frac{C_P}{C_S}\right]^{-2}\frac{dC_S}{dt}.$$

In Schottky's approximation:

$$\frac{dC}{dt} = \frac{d}{dt}\left(\frac{2\epsilon\epsilon_0 eN}{Q_S}\right) = -\frac{2\epsilon\epsilon_0 eN}{Q_S^2}\cdot\frac{dQ_S}{dt}$$
$$= -\frac{1}{V_S}\cdot\frac{dQ_S}{dt},$$

where V_S is the voltage across the capacity C_S. Therefore,

$$\frac{dQ_0}{dt} = -\frac{dQ_S}{dt}\left\{1+\frac{V}{V_S}\left(\frac{C_P}{C_S}\right)^2\left[1+\frac{C_P}{C_S}\right]^{-2}\right\}.$$

Since

$$\frac{V}{V_S} = \frac{V_P+V_S}{V_S} = 1+\frac{V_P}{V_S} = 1+\frac{C_S}{C_P}$$

one finally finds

$$\frac{dQ_0}{dt} = -\frac{dQ_S}{dt}\cdot\frac{1+2\,C_P/C_S}{1+C_P/C_S}. \qquad (2)$$

This result indicates that under restricted experimental conditions the change in surface conductance represents the change in surface charge. In the experiments described later, the condition $C_P \ll C_S$ is fulfilled. Therefore:

$$\frac{dQ_0}{dt} \approx -\frac{dQ_S}{dt}.$$

The technique for investigating the transient behaviour of surface states consists of a sudden change in the surface concentration of majority carriers by applying a voltage pulse to the field plate. The polarity is such that the carriers are pushed away from the surface. The reduction in the concentration of majority carriers causes the surface states to readjust their occupancy. If the reduction is strong enough a generation process only will occur and the recombination process can be neglected. This simplifies the analysis.

In order to evaluate the transient behaviour of the surface conductance, one has to connect the measured relaxation time τ with the cross section of the reacting state. Applying the principle of detailed balance,[2] one finds for the reaction of a state with the conduction band:

$$\sigma_n(T) = (\tau\gamma n_1 v_T)^{-1} \qquad (3)$$

and for a reaction with the valence band

$$\sigma_p(T) = (\tau p_1 v_T/\gamma)^{-1} \qquad (4)$$

for both donor- and acceptor-like states. γ is the statistical factor of the level; n_1 and p_1 stand for the surface concentration of electrons and holes respectively, if the Fermi level coincides with the level in question; v_T is the thermal velocity of the majority carriers.

A similar analysis can be carried out for the case that more than one surface state contributes to the relaxation process. If they differ in their energy position, one can separate them, in general, by a study of the temperature dependence of the relaxation process.

EXPERIMENTAL PROCEDURE AND RESULTS

The silicon samples were cut in the form of rectangular solids of size 1 cm by 0·3 cm by 0·05 cm. Gold wires were bonded to the sample ends and care taken that the contacts were ohmic. After polishing, the surfaces were etched with CP-4, or white etch, a mixture of nitric and hydrofluoric acid, to remove surface damage. The resistivity of the samples varied between 10 and 200 Ω-cm. Part of the material was grown in standard facilities The other part was floating-zone refined Merck material.

The samples, which were held at a vacuum of the order of 10^{-6} mm mercury during part of the measurement, were mounted in a chamber and connected through a sapphire to a temperature bath on one side. A metalized sapphire, which provided electrical isolation and insured thermal equilibrium, was mounted on the other side as a field plate and was spaced with a mylar sheet of 1 mm thickness as a dielectric. The choice of this thickness was influenced by consideration of equation (2) in order to make $C_P/C_S \ll 1$ and the

desire to keep the electric field sufficiently uniform over the surface area in spite of surface roughness. The voltage pulses applied to the field plate had a rise time of about 20 μ sec and a negligible drop of the pulse height. A constant current source provided a current of about 0·1–3 mA through the sample. The ends of the samples were connected to the inputs of a differential amplifier. By keeping the impedance of this circuit of the order of 10 kΩ, the feed-through signal from the field plate could be cancelled easily. A switch provided the possibility of reversing the current through the sample in order to check that the observed signal was not influenced or caused by rectification at the contacts. The change in surface conductivity could be observed directly with a Tektronix scope 532 using the d.c. channel of the high gain differential preamplifier. With this instrumentation, relaxation phenomena between 60 μ sec and 0·1 sec could be measured. For longer time constants, the signal was fed into a Brush recorder which covered the range between 10^{-2} and 3 sec. For still longer time constants, a Leeds and Northrup recorder was used, covering the range from 2 sec up. It turned out that the results in the overlapping ranges for the different instrumentations were identical. The pulse height for this series of experiments varied between 300 and 1500 V. Occasionally, voltages up to 3000 V were used.

During the run, the voltage was kept constant and only the temperature was changed slowly and continuously. The time constants were either read directly from the screen of the Tektronix scope or evaluated from photographs or recorder traces. The runs were taken with increasing and decreasing temperature, and were repeated with different applied voltages.

An evaluation of the τ values is possible if one decides to set the unknown statistical factor γ equal to unity and if one makes the further assumption of a temperature-independent cross section. Equations (3) and (4) can then be simplified to

$$\sigma_n \exp\left[-(E_C - E_{A,D})/kT\right] = (\tau N_C v_T)^{-1} \quad (3^*)$$

and

$$\sigma_p \exp\left[(E_V - E_{A,D})/kT\right] = (\tau N_V v_T)^{-1} \quad (4^*)$$

An approximate level position at $T = 0$ and an effective cross section can therefore be obtained from a plot of log $(\tau N_C v_T)^{-1}$ and log $(\tau N_V v_T)^{-1}$ against $1000/T$. The results of runs with two n-type samples are shown in Fig. 1. M2–2A is zone-refined Merck material, K77–1A is grown in standard facilities. The obtained results show a discrete level at 0·52 eV below the conduction band and an effective cross section for electrons of 3×10^{-16} cm^2.

In Fig. 2 the result obtained with three p-type samples are shown. Two of the samples are cut from a silicon crystal grown in standard facilities. The third sample is zone-refined Merck material. The activation energies vary between 0·52 eV and 0·56 eV indicating a discrete level about 0·54 eV above the valence band. The cross section for the capture of holes is about 2×10^{-15} cm^2.

Two additional experiments with gaseous ambients were carried out. An n-type sample, 4M–6A, was treated with ozone. The results of two runs are shown in Fig. 3. They yield an activation energy of 0·63 eV and a capture cross section for electrons of the order of 10^{-13} cm^2.

After the runs in vacuum, described in Fig. 2, the p-type sample 206–1–1A, was exposed to dry ammonia. The results of this pulsed field effect run are shown in Fig. 4. The activation energy is now 0·72 eV and the capture cross section for holes has become of the order of 10^{-12} cm^2.

CONCLUSIONS

It should be emphasized that the quoted values for the cross sections indicate only an order of magnitude. This uncertainty stems from the unknown statistical factor for these levels and the unknown temperature dependence of the cross section. The influence of these approximations, however, upon the determination of the level position is not significant.

A cross section of the order of 10^{-15}–10^{-16} cm^2 is indicative of a neutral center. This suggests that, in n-type material the observed level is acceptor-like; in p-type silicon, donor-like. This is in agreement with the observation that the Fermi level in silicon surfaces appears to be locked at midgap.

The question arises whether the two midgap levels are present simultaneously. The observation in standard field effect experiments tends to give evidence that both levels are present. The puzzling fact that it is practically impossible to produce an inversion layer by application of a strong field (unless one goes to elevated temperatures) can be

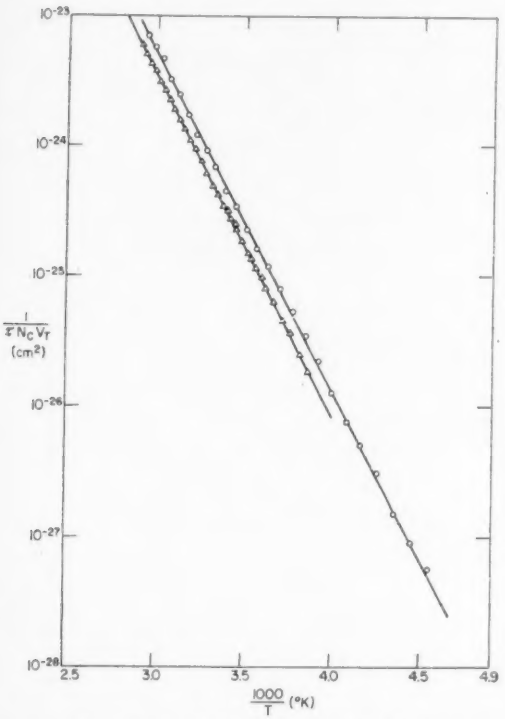


FIG. 1. Pulsed field effect results for two different n-type silicon samples.

○ M2–2A
n-type, floating zone
$\epsilon = 0{\cdot}52$ eV.
$\sigma_n = 3{\cdot}9 \times 10^{-16}$ cm²
25 Ω-cm (phosphorus)

△ K77–1A
n-type
$\epsilon = 0{\cdot}52$ eV.
$\sigma_n = 2{\cdot}3 \times 10^{-16}$ cm³

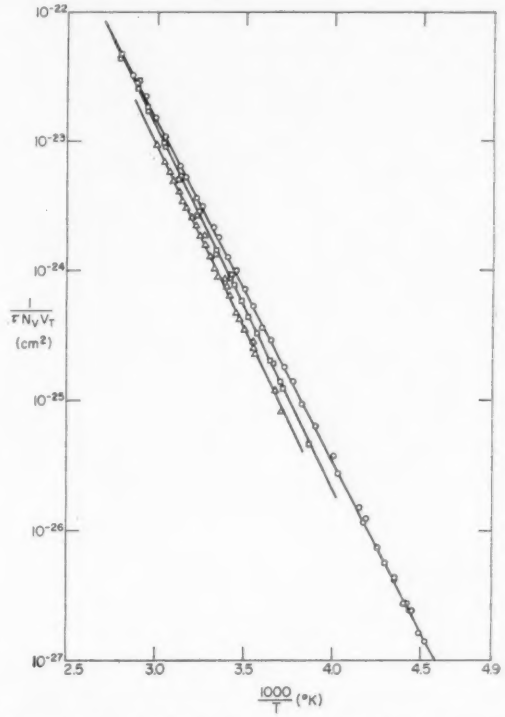


FIG. 2. Pulsed field effect results for three different p-type silicon samples.

○ K121–2
p-type
$\epsilon = 0{\cdot}52$ eV
$\sigma_p = 1{\cdot}1 \times 10^{-15}$ cm²
60 Ω-cm (boron)

□ K121–3C
p-type
$\epsilon = 0{\cdot}55$ eV
$\sigma_p = 2{\cdot}5 \times 10^{-5}$ cm²
60 Ω-cm (boron)

△ 206–1–1A
p-type, floating zone
$\epsilon = 0{\cdot}56$ eV
$\sigma_p = 2{\cdot}8 \times 10^{-5}$ cm²
21 Ω-cm (boron)

explained with this assumption. The combined effect of these two levels leads to a rigid locking of the Fermi level at midgap, which one level alone would not do. The same line of argument explains the observation that an introduction of light removes the observed non-equilibrium loop in field effect studies.

A clear cut answer, however, to this question can be given through experiments with ambients. The Merck n-type sample 4M–6A was exposed to ozone which is known to pull the surface towards the p-side. Since the activation energy of 0·63 eV, as shown in Fig. 3, together with $(E_D - E_V) = 0{\cdot}54$ eV adds up to approximately the full bandgap at the temperature of absolute zero, one can conclude that the observed relaxation effect is due to the same donor-like level which was constantly observed in p-type material. This is strongly supported by the magnitude of the cross section which is characteristic for an attractive center in agreement with the previous interpretation. In order to show that both levels are present simultaneously also in p-type silicon, sample 206–1–1A was exposed to dry ammonia as an ambient. Ammonia is known to pull the surface towards the n-side. The observed activation energy of 0·72 eV as shown in Fig. 4 together with $(E_C - E_A) = 0{\cdot}52$ eV adds up to approximately the full bandgap. Therefore, the argument used above leads to the conclusion that in p-type material the same

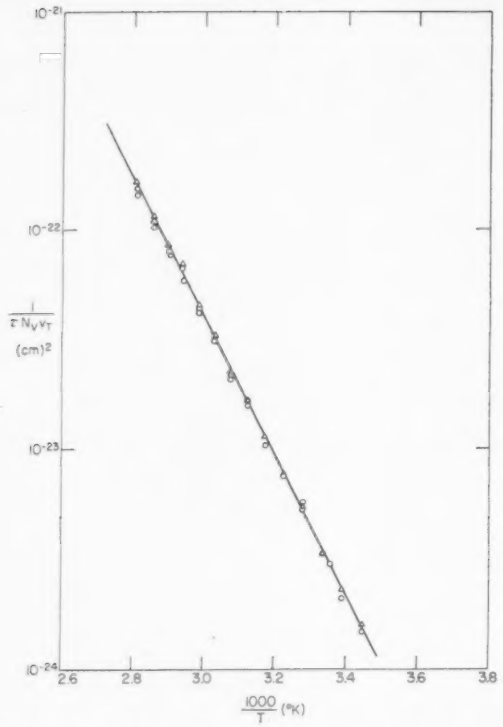


FIG. 3. Pulsed field effect results for an *n*-type silicon sample under the influence of ozone as an ambient.
4M–6A *n*-type silicon 300 Ω-cm.
$\epsilon = 0{\cdot}63$ eV. $\sigma_n = 1{\cdot}3 \times 10^{-13}$ cm².
○ = Run No. 1. △ = Run No. 2.

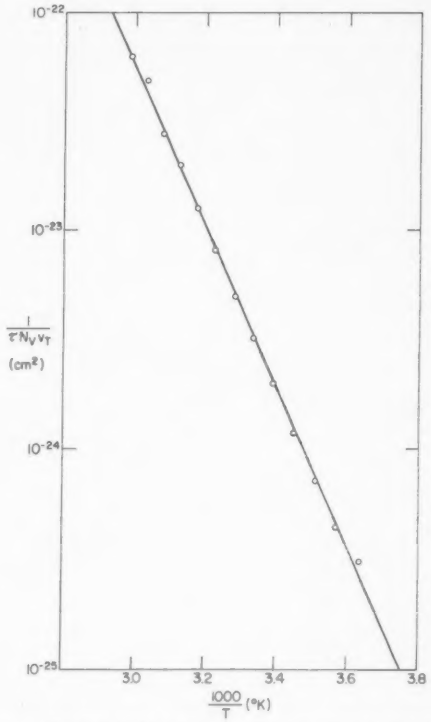


FIG. 4. Pulsed field effect results for one of the *p*-type silicon samples under the influence of ammonia as an ambient.
206–1–1A *p*-type Merck.
$\epsilon = 0{\cdot}72$ eV
$\sigma_p = 4 \times 10^{-12}$ cm²

acceptor-like level is also present which was always observed in *n*-type material. The cross section of this level supports the previous interpretation of its charge character.

Another question should be asked. Do bulk levels at midgap contribute to the observed relaxation phenomenon? If bulk states contribute to the relaxation process, an additional study must be carried out. The voltage dependence of $\Delta\Sigma(t,V)$, the change in surface conductance, in general gives this information. This has been demonstrated in gold-doped germanium.[3] The leading idea is that an increase in pulse height finally will empty the entire charge out of the surface state and a further increase has no effect on the quantity $\delta(\Delta\Sigma) = \Delta\Sigma(0,V) - \Delta\Sigma(00,V)$, which becomes constant. If the relaxation effect is caused by a bulk-level, however, $\delta(\Delta\Sigma)$ will still increase with increasing pulse height, and will not reach saturation. The reason is that increasing pulse height still widens the space charge layer and therefore still further increases the number of reacting bulk states. The criterion that $\delta(\Delta\Sigma)$ saturates with increasing pulse height does not work too well in this case. The high density of the surface states prevents $\delta(\Delta\Sigma)$ from reaching a saturation. Although this criterion cannot be used here to prove that only surface states are present, an extrapolation of the onset of a saturation effect leads to a rough estimate of the surface density and yields the order of 10^{13} states/cm².

As ordinarily-grown silicon contains an appreciable concentration of oxygen in the lattice, one could suspect that the oxygen would cause the observed midgap levels. Therefore, the experiments were carried out also with *n*-type and

p-type material from Merck, with an oxygen concentration of the order or less than $10^{16}/cm^3$. The results which can be seen by comparing the data in Fig. 1 and Fig. 2 were the same as in ordinary silicon. Hall data at elevated temperatures also did not indicate the presence of bulk states close to midgap though they are hard to detect.

The results of the application of ambients again answers the question above. If one assumes that the states in question are bulk states, one cannot explain that the application of ambients causes a change in the relaxation process.

The two questions stated above are, therefore, answered at the same time. The experiments with the ambients show that both levels are present simultaneously and that the observed relaxation processes are predominantly due to surface states, because ambients could not have changed the transient behaviour of bulk states.

The knowledge of the predominant surface states in silicon was an important part of the investigation of bulk states by the pulsed field effect method. The measurement of the position and the cross section of bulk levels introduced by neutron bombardment could be taken easily because of the particular properties of the surface states in silicon.[4–6] It should be added that no change in the surface states due to neutron bombardment could be observed for the radiation doses used ($\approx 10^{14}/cm^2$).

Acknowledgements—The author wishes to express his gratitude to P. SMAKULA for his help and painstaking care in taking the data, to Dr. S. R. ARNOLD for his support in the design of the measuring circuit, and to L. RUBIN for help in the instrumentation.

REFERENCES

1. STATZ H., DEMARS G., DAVIS L. Jr. and ADAMS A. Jr., *Phys. Rev.* **106,** 455 (1957).
   STATZ H., DEMARS G., DAVIS L. Jr. and ADAMS A. Jr., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 139. University of Pennsylvania Press, Philadelphia (1957).
2. RUPPRECHT G., *Phys. Rev.* **111,** 75 (1958).
   A more detailed calculation will be given in a forthcoming paper.
3. RUPPRECHT G., *Bull. Amer. Phys. Soc.* **3,** 377 (1958).
4. KLEIN C. A., *J. Appl. Phys.* **30,** 1222–1231 (1959).
5. RUPPRECHT G. and KLEIN C. A., *Phys. Rev.* **116,** 342 (1959).
6. KLEIN C. A. and RUPPRECHT G., to be published.

VOL. 14 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 214–219. Printed in Great Britain.

THE EVALUATION OF GERMANIUM SURFACE TREATMENTS

S. R. MORRISON

Honeywell Research Center, Hopkins, Minnesota

Abstract—Studies have been made of the variations of electrical surface properties with surface treatment. The variables measured are the field effect mobility and the surface recombination velocity. A report is presented of the results with various etch systems, the influence of contamination in the etch or rinse water (some contaminants are found seriously to alter the characteristics with as low as 0·5 p.p.m.) and the variation of surface properties during heat treatment. A method has been developed for rapid estimation of surface barrier, density of fast states and reproducibility of the surface (with the large number of variables associated with surface treatments, rapidity in accumulation and interpretation of data is essential). These parameters (surface barrier, etc.) are obtained from the measurements discussed above and a set of calibration curves developed for the purpose. The method requires the energy levels of the fast states to be independent of surface treatment (or another set of calibration curves be developed for the new set of energies) but is self-checking with respect to this energy level parameter. Of those treatments used to date, the only one which has not been comparable in energy levels produced is the electrolytically etched surface, and a more exact analysis by WALLIS and WANG[4] has indicated also that here the energy levels are at considerably different values. The experimental results are interpreted in terms of fast state density and surface barrier using this method. Evidence will be presented that under certain treatments water vapor can change the density of recombination centers.

INTRODUCTION

THE influence of chemical surface treatments on the physical properties of germanium has received a great deal of attention in recent years, but is such a complex problem that it is still poorly understood. The problem has been approached in the past from two directions. On the one hand, there has been the principally chemical approach, as that of FAUST[1], CRETELLA and GATOS[2], and ELLIS[3], where a wide variety of chemicals is applied to the surface, and the action is reported primarily in terms of the appearance or of various chemical properties of the surface. On the other hand, there has been the principally physical approach used by WALLIS and WANG[4], HARNIK and MARGONINSKI[5], and others as reviewed by MANY[6]. Here, although careful measurements provide numerical values for the energy and density of fast states following surface treatments, the experimental measurements are very difficult and slow. Thus it is impractical to provide the wide variety of treatments desired for an extensive study relating variations in chemical treatment to variations in physical properties.

The present study was undertaken to provide a compromise between the cited approaches; to investigate a wide variety of chemicals, and yet obtain information regarding the resulting electrical behavior. With this purpose in mind, it was necessary to compromise somewhat on the exactitude of the method used by WALLIS *et al.* in order to allow rapid accumulation of data.

METHOD

The method adopted for this study utilizes simple field effect mobility, μ_{FE}, and surface recombination velocity, s, measurements, each of which can be made in a few minutes, together with a calibration curve to be discussed below. The field effect measurement is the usual low field a.c. measurement; the value of s is obtained from a photoconductive lifetime measurement. The germanium used was between 5 and 5·5 Ω-cm, with bulk lifetime greater than 100 μ sec.

The use of a calibration curve enables the s and μ_{FE} measurements to be interpreted visually in terms of the density of fast-states and the surface

barrier. The calibration curves are derived experimentally, but the theoretical basis is of interest, and will be discussed first.

We will investigate the theoretical reasoning behind the calibration curves with the help of Fig. 1. No attempt is made at rigor in this illustration.

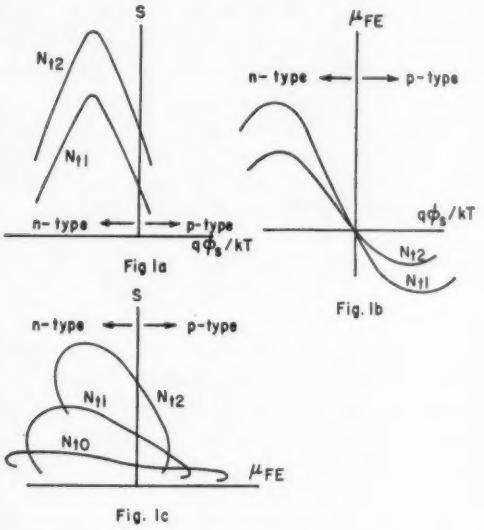


FIG. 1. Theoretical basis for calibration curves. $-N_t$ represents the density of fast states:

$$N_{t2} > N_{t1} > N_{t0}.$$

Several workers, as reviewed by MANY[6], have shown that in a simple case the relation between s and the surface potential ϕ_s (where ϕ_s is the difference between the Fermi level and the intrinsic Fermi level at the surface) should be similar to Fig. 1(a). At constant ϕ_s, s is proportional to N_t, the density of fast states. The relationship[7] of μ_{FE} to ϕ_s is sketched in Fig. 1(b), where the fact is indicated that at constant ϕ_s, μ_{FE} decreases, as N_t is increased. Combining these curves graphically to eliminate ϕ_s, we observe that s as a function of μ_{FE} should follow a set of curves such as shown in Fig. 1(c), which again are curves of constant N_t.

Although it is thus easily shown theoretically that curves of the form of Fig 1(c) should be characteristic of surfaces, the theory is not sufficiently advanced to derive the curves rigorously. The approach we have used is to develop the curves experimentally. Samples were prepared with various surface characteristics, and were cycled between dry oxygen and wet (20 per cent RH) nitrogen, while s and μ_{FE} were followed. The assumption was made that with this modest treatment no change in N_t would occur. About two hundred such measurements were made,* covering most of the range of s and μ_{FE} readily available. In Fig 2 a few of these measurements are shown

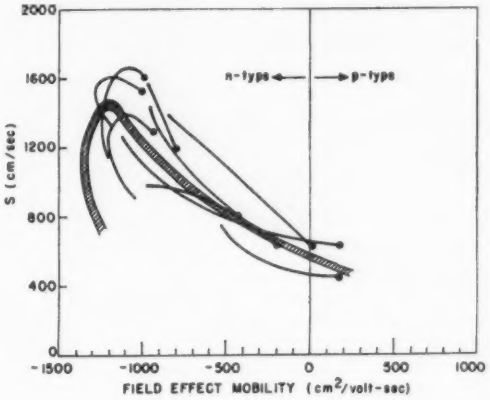


FIG. 2. Development of calibration curves from experimental data.

as a typical illustration. The dots show the measurement in dry oxygen. The curves are the path the parameters take as the surface becomes more n-type with wet nitrogen. With data of this type available, the average is taken by inspection, and the curves corresponding to Fig 1(c) are determined. The broad curve of Fig 2 is the curve obtained from the experimental curves shown, taking into consideration the numerous curves not recorded here.

Figure 3 shows the set of calibration curves developed from the data available. The interpretation follows directly from the theoretical basis discussed in conjunction with Fig. 1. Each of the individual curves of Fig. 3 is thus assumed to represent a constant fast state density. This density (in arbitrary units) can be roughly estimated as follows. From the theory of recombination[8, 9] at a constant barrier height, s is proportional to N_t, the

* Surface treatments used in developing these curves were varying mixtures of HF–HNO_3, with and without copper nitrate addition in the etch or rinse water, and with and without heat treatment. A few CP-4 treatments were also included.

density of fast states. In Fig. 3 the barrier height is constant at $\mu_{FE} = 0$, so the density of fast states for each of the curves of Fig. 3 should be proportional to the value of s at $\mu_{FE} = 0$. This particular interpretation requires a dominant state at this surface barrier, and the density obtained is the density for this dominant state.

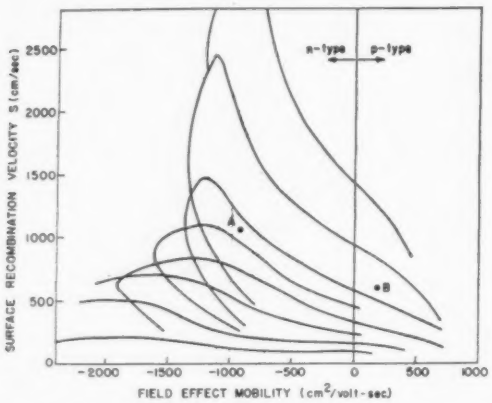


FIG. 3. "Calibration curves"—Experimentally determined variation of surface recombination velocity with field effect mobility at constant fast state density.

The use of Fig. 3 in evaluating new surface treatments is as follows. In principle μ_{FE} and s (and if necessary $ds/d\mu_{FE}$, as obtained from an ambient change) provide all the necessary information. However, with the calibration curves available, it is much easier to compare the characteristics of two different surface treatments. If, for example, two surface treatments yield parameters represented by points A and B (Fig. 3), one can say by inspection that point A has a lower density of fast states than point B, despite the fact that point A has a higher s. In fact, depending on the confidence one has that the assumptions below are valid, one can even become semi-quantitative and relate the density of fast states numerically, through the use of s at $\mu_{FE} = 0$, and relate the surface barriers semi-quantitatively by relating it to the distance of the point between $\mu_{FE} = 0$ and $\mu_{FE} = \max$.

RESTRICTIONS AND ASSUMPTIONS

The first restriction on the use of the calibration curves is the requirement that the temperature and resistivity of the sample be the same as those from which the calibration curves were obtained. Together with these obvious requirements, however, there are more subtle restrictions arising from implicit assumptions made in the derivation of the calibration curves.

In order to use the calibration curves even qualitatively, data must fit the curves. That is, with any surface treatment, a measurement must be made of the slope of the curve, $ds/d\mu_{FE}$, as the ambient changes, and if this does not correspond to that of the calibration curve, one cannot interpret the results using the calibration curve. At a given s and μ_{FE}, in general $ds/d\mu_{FE}$ will correspond only if (a) the energy levels, (b) the *relative* densities (if more than one state is important) and (c) the cross sections of the fast states are the same as those represented by the calibration curves. Conversely, if $ds/d\mu_{FE}$ does not correspond, then (a), (b) and (c) are not the same as those represented by the calibration curves. In our experiments, to be discussed below, the only case investigated where the surface treatment did not meet this requirement was the case of electrolytically etched surfaces. WALLIS and WANG[4] also find that in this case the fast state energy level is quite different from the norm.

There is one implicit assumption in the above discussion, that is that (a), (b) and (c) above, are valid throughout the treatments used in developing the calibration curves. The correctness of the assumption is indicated by (i) the mutual consistency of the results used to produce the calibration curves; (ii) the results of WALLIS and WANG[4] who find that these characteristics are indeed approximately independent of surface treatment for most common surface treatments. If the assumption fails to a certain degree, then the use of the calibration curve becomes inaccurate to that degree.

RESULTS AS A FUNCTION OF ETCH TREATMENT

Three empirical rules seem to evolve from our investigations.

(i) The most significant result is the demonstration that very small amounts of impurity in the etch or rinse water have a controlling effect on the resulting surface characteristics. The most important of these impurities are copper, antimony and silver.

The determining characteristic may be the position relative to germanium in the electromotive series.

(ii) If the etch and rinse water are free of such ions, the surface is always *n*-type. The higher the purity, the more *n*-type, apparently, although our sensitivity obviously limits the extension of this rule very far.

(iii) Another empirical rule which seems to be followed is that the higher the per cent concentration of HNO_3 in the etch, the more *p*-type the surface, either because it is more sensitive to impurities or due to inherent properties of the pure etch.

Figure 4 shows typical results with various etch treatments followed by a rinse in relatively pure (7 MΩ-cm) water. These readings, as with all measurements to follow, were made at room temperature in dry oxygen, with each experiment point an average of at least four runs. The probable error estimated from these four runs is shown. The dots are added to indicate the most *p*-type point in the range (to establish, where doubt exists, on which branch of the curve the characteristics lie) and do not necessarily represent a measured point.

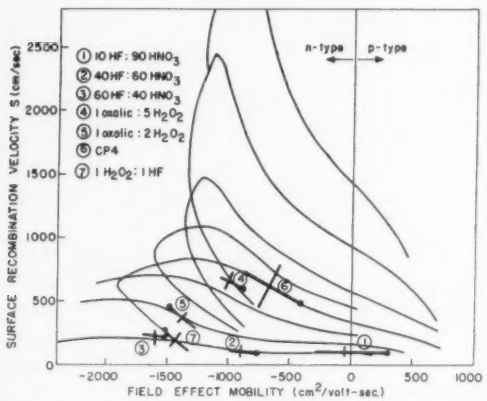


FIG. 4. Surface parameters for some common etch treatments.

Points 1, 2 and 3 represent HF–HNO_3 mixtures with increasing HF concentrations. That the 90–HNO_3 result is about intrinsic corresponds to rule (iii) above. The results with CP-4 are observed to show a high density of recombination centers and very poor reproducibility. Adding impurities can make *s* lower by making the surface more *p*-type, but the reproducibility is still poor. The source of the difficulty seems to be the bromine. With the bromine deleted, CP-4 acts just like an HF–HNO_3 etch.

Figure 5 shows the influence of impurities. Measurements are shown of the field effect mobility as a function of the amount of copper nitrate added to the rinse water following an HF–HNO_3 etch. This relationship effectively demonstrates the evidence for the three rules above. The two points at about 0·04 p.p.m. copper nitrate have vertical probable error lines associated with them. These points represent results with a "pure" water rinse, with no copper intentionally added, and the vertical lines are estimated limits for the copper concentration, as estimated from water resistivity measurements. The value of *s* for these measurements are all between 100 and 200 cm/sec. The influence of copper ions in sending the surface *p*-type is obvious, as is the

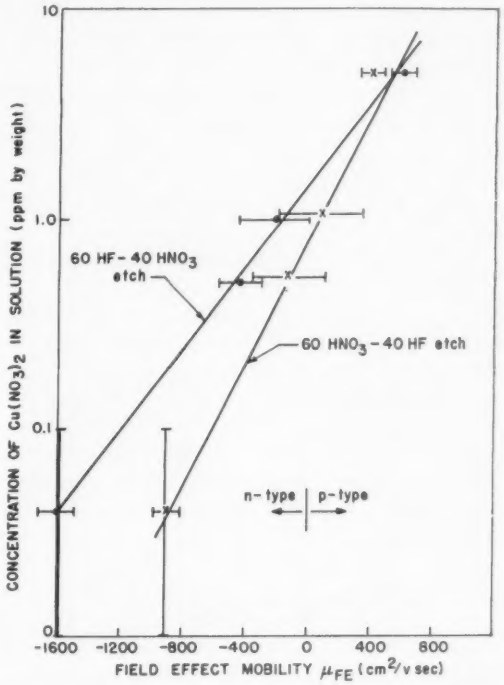


FIG. 5. The influence of copper ions in the rinse water, making the surface *p*-type.

observation that the surface with a "pure" chemical treatment would undoubtedly be more n-type than those recorded here, there being no evidence of saturation.

Similar results are observed if the impurity is added to the etch itself rather than to the rinse water. Here, however, the concentration necessary is much larger, by a factor of the order of 100, in order to exert the same influence. It is possible that the influence may arise from residual copper nitrate being carried to the rinse water when the sample is transferred. Etches such as the Westinghouse silver etch, the copper etch or antimony etch, with 4:2:4 of HF:HNO_3:H_2O plus 0·02 g/cm^3 of the respective metal salt, inevitably result in a p-type surface ($s = 200 \pm 100$ cm/sec, $\mu_{FE} = 300 \pm 100$ cm^2/V sec.

We have performed measurements designed to produce a uniform, low s, p-type surface treatment and have developed the following as most satisfactory. The sample is first etched with 70 HNO_3:30 HF, and rinsed in high resistivity water (>5 MΩ-cm); the second rinse is water plus 100 p.p.m. $Cu(MO_3)_2$, and a third rinse in high resistivity water to remove excess copper is provided before air drying. This etch results in a p-type surface ($s = 40 \pm 20$ cm/g, $\mu_{FE} = 200 \pm 50$ cm^2/V sec).

RESULTS AS A FUNCTION OF HEAT TREATMENT

In Fig. 6 is shown the variation of the surface parameters as a function of time and temperature

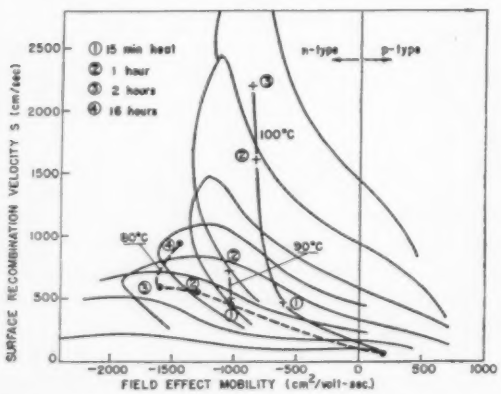


FIG. 6. The variation of surface parameters with dry heat treatment.

of dry oxygen heat treatment. The surface is initially etched to produce a p-type surface. With short heat treatment it becomes more n-type, with little change in density of fast states; with further treatment the density of fast states increases, with the surface remaining at about the same barrier height.

The relative humidity during the heat treatment is of interest. Fig. 7 shows the variation with dry

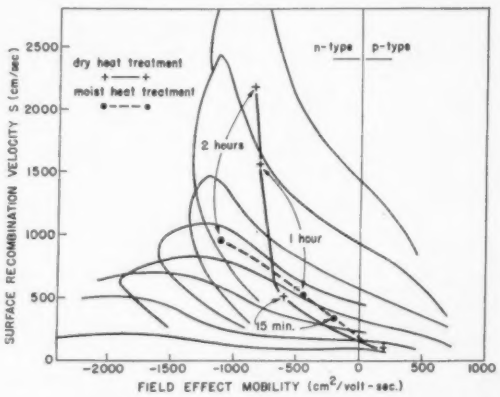


FIG. 7. The influence of small amounts of moisture on the variation during heat treatment.

heat treatment and moist heat treatment at 100°C. The moisture content for the moist treatment corresponds to 100 per cent RH at room temperature (about 7 mm H_2O). For the dry heat treatment the oxygen is desicated at room temperature by Drierite.

With larger amounts of moisture the direction can be actually reversed. Fig. 8 shows the change in parameters using 100 per cent RH at various temperatures with a one-hour treatment. Here again, of course, the measurements are made at room temperature in dry oxygen. Prior to the humid treatments, the samples had been baked to establish the initial high s condition. At 40°C, the fast states are apparently unaffected, although the surface goes more p-type. At 60 or 70°C, the fast state density apparently decreases considerably with the humid treatment. At 90°C, however, the fast state density is again relatively unaffected. This last result may be due to the fact that at 90°C we are in the temperature range where dry heat treatments increase the fast state density, so there is a conflict between the two effects.

From these results, it appears that dry heat treatment tends to make the surface more *n*-type and to increase the density of fast states. Humid heat treatment tends to make the surface more *p*-type and to decrease the density of fast states.

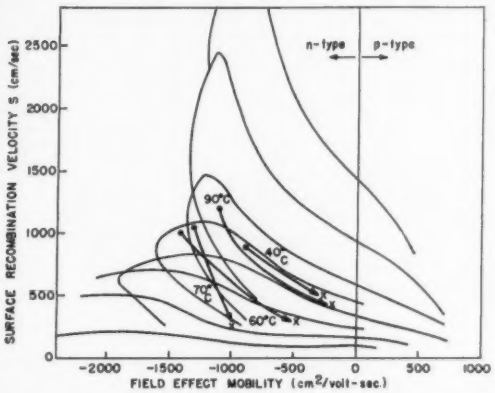


FIG. 8. Humid heat treatment variations.

The temperature range where dry heat treatment effects are rapid is 80°C up, that for humid is about 50°C up.

A possible interpretation of these observations is that water vapor is dissolved in the oxide layer[3] and reacts with the fast states, possibly changing their energy level to an ineffective position. Dry heat treatment bakes off the water vapor, regenerating active fast states, and humid heat treatment encourages the transfer of water vapor back through the oxide, deactivating the fast states.

A simple and significant phenomenological interpretation of the experimental results may be that there exists an equilibrium density of recombination centers and an equilibrium surface barrier at each temperature and relative humidity. A change in either of these variables will cause a drift toward a new equilibrium state. If valid, such an interpretation should be extremely valuable in explaining some of the long time drifts of surface properties, such as are observed with transistors.

Acknowledgements—The author wishes to express his indebtedness to R. BILLETE, R. HAGER and D. ZANDER, who assisted with the experimental measurements.

REFERENCES

1. FAUST J. W. Jr., *Trans. Electrochem. Soc.* **106,** 201c (abstr.) (1959).
2. CRETELLA M. C. and GATOS H. C., *Trans. Electrochem. Soc.* **105,** 487 (1958).
3. ELLIS S. G., *J. Appl. Phys.* **28,** 1262 (1957).
4. WALLIS G. and WANG S., *Trans. Electrochem. Soc.* **106,** 231 (1959).
5. HARNIK E. and MARGONINSKY Y., *J. Phys. Chem. Solids* **8,** 96 (1959).
6. MANY A., *J. Phys. Chem. Solids* **8,** 87 (1959).
7. SCHRIEFFER J. R., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 55. University of Pennsylvania Press, Philadelphia (1957).
8. STEVENSON D. T. and KEYES R. J., *Physica, s'Grav.* **20,** 1041 (1954).
9. MANY A., HARNIK E. and MARGONINSKY Y., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 85. University of Pennsylvania Press, Philadelphia (1957).

VOL. 14 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 220–225. Printed in Great Britain.

THE SURFACE RECOMBINATION ON SILICON CONTACTING AN ELECTROLYTE

H. U. HARTEN

Philips Laboratories, Hamburg, Germany

Abstract—(1) Excess carriers recombine at the silicon–electrolyte interface in a way not very different from that at dry surfaces. (2) The recombination velocity depends on the surface potential in a bell-shaped manner similar to the "Stevenson–Keyes curve". (3) The surface potential can be varied in a relatively wide range by applying only some tenths of a volt between silicon and electrolyte. (4) A part of the applied voltage drops across the oxide layer present at the silicon surface. This part is larger the thicker the oxide layer in relation to the space charge layer inside the silicon.

INTRODUCTION

EXTENDED work has been done by many authors on metallic electrodes in electrolytes. Besides chemical methods mostly the current–voltage characteristic and the capacitance–voltage characteristic have been observed. These methods also lend themselves for investigation of semiconducting electrodes and some work has been done in this direction too.[1] In addition, information may be obtained about effects present in semiconductors only, but missing in metals. One of these effects is the surface recombination. Some observations on this subject are here presented.

METHOD

The recombination velocity of excess carriers at the surface of a semiconductor can be determined from photovoltaic measurements in an arrangement as shown in Fig. 1: A thin semiconducting wafer with p–n junction at one surface is illuminated at the opposite surface. If nonpenetrating light is used, excess carriers are generated just below the surface and the more of them recombine at the surface the fewer will reach the junction. To a rough approximation, the photo voltage at the junction will thus be inversely proportional to the surface recombination velocity. If penetrating light is used, the carriers are generated more homogeneously in the bulk and the surface has very little influence. The correlation between the photo-voltaic sensitivity and the depth of penetration of monochromatic light has been calculated theoretically[2] and a fairly good agreement with experimental results has been found.[2, 3] Therefore, vice versa, the surface recombination velocity s can be determined from such measurements (as well as the distance d between surface and junction and the diffusion length L in the bulk, homogeneous distribution assumed).

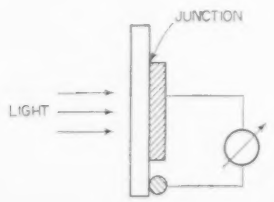


FIG. 1. Principle of the arrangement used for measuring the surface recombination velocity.

Results similar to the ones with dry surfaces can be obtained with surfaces contacting an electrolyte, at least with silicon. In Fig. 2 the experimental arrangement is shown schematically. The silicon wafer is about 0·5 mm thick; the junction and ohmic contact are made by an alloying technique. The light, coming from a high-pressure mercury lamp energized by the a.c. mains, is modulated nearly sinusoidally at 100 c/s, thus the photo voltage can be measured by a tuned wave analyser. The light has passed through a single prism monochromator. The potential between the silicon and

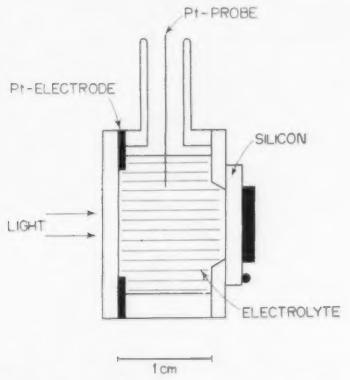


FIG. 2. Mounting of the electrolytic cell.

VOL. 14 1960

the electrolyte can be measured at the platinum probe with a tube voltmeter and can be varied by applying a d.c. voltage at the platinum electrode.

BASIC OBSERVATION

In Fig. 3 three typical sets of measurements are shown together with the calculated values (solid lines). The s-scale is related to the horizontal parts of the curves. The agreement between theory and experiment is found to be as good as in the case of real surface recombination at a dry silicon surface; even the disagreement in the lowest curve has been observed in a similar manner at dry surfaces. A satisfactory explanation cannot be given at present. The electrolyte used in these measurements was an aqueous solution of Na_2SO_4 (ca. 1wt.%) but similar results have been also obtained in distilled water and in tetraethylammoniumiodide. Non-aqueous solutions were not investigated.

The above observations suggest the assumption that there is no fundamental difference between the recombination processes at a dry silicon surface and at a surface in contact with an electrolyte.

THE "STEVENSON-KEYES-CURVE"

The three curves of Fig. 3 are obtained with a single silicon wafer varying the potential difference between silicon and electrolyte only by applying a voltage to the platinum electrode. The potential difference is measured at a platinum wire acting as a potential probe. The fact that the contact potential between the wire and the electrolyte was fairly constant under conditions present in the recombination experiments has been checked by means of a calomel electrode in a special arrangement.

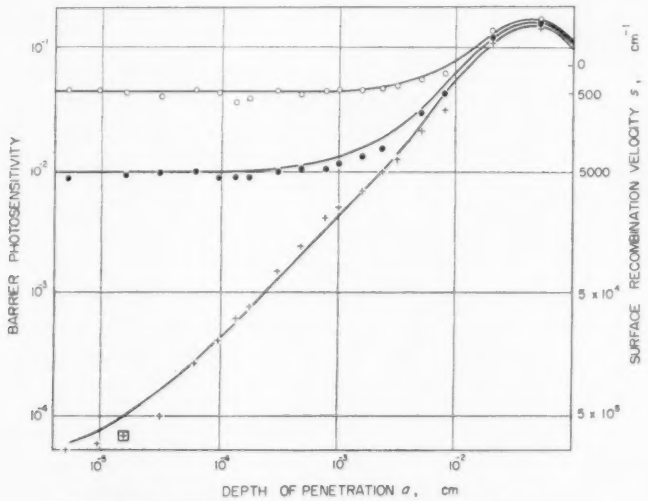


FIG. 3. Sensitivity of the barrier photo effect vs. depth of penetration of light. The solid lines are calculated with the formula of Ref. (2). The s-scale is related to the horizontal parts of these curves. The measuring points are taken of three different voltages V_a, namely $+0{\cdot}5$ V, $-0{\cdot}3$ V, $+0{\cdot}1$ V (from top to bottom).

Thus the measured voltage V_a between wire and silicon equals the potential difference V between electrolyte and silicon with the exception of an additive term which, however, is not relevant even if it varies between different sets of measurements.

It follows from Fig. 3 that the recombination velocity s at the silicon electrolyte interface can be altered to a large extent by varying the voltage V_a within some tenths of a volt only. In order to obtain the complete dependence it is sufficient to measure the sensitivity at a single (and small) depth of penetration of light. Knowing d and L from Fig. 3, the recombination velocity s can be determined merely from this sensitivity. L is the diffusion length and d the distance from surface to junction.

A typical result is given in Fig. 4 with the sensitivity scale on the left hand and the calculated

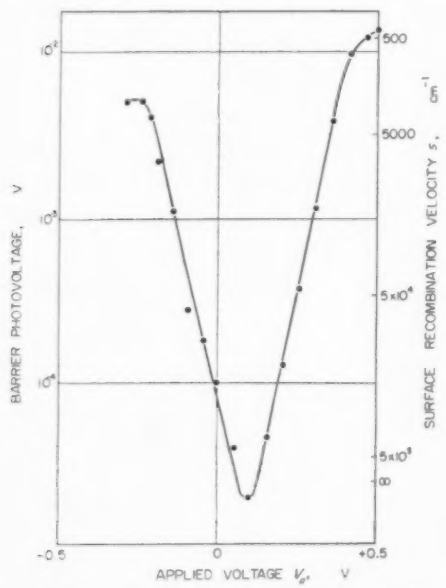


FIG. 4. Barrier photo voltage vs. the voltage V_a applied between silicon and electrolyte (depth of penetration of light 0·25μ).

s-scale on the right (violet mercury line, wave length 0·40 μ, depth of penetration 0·25 μ). Values of s beyond infinity are simulated within a small region; this stems from the disagreement mentioned above in the lowest curve of Fig. 3.

The V_a-region within which s can be measured is limited by the current in the electrolytic circuit. As shown in Fig. 5 for n-type silicon this current has relatively small values (<1 μA/cm^2) if V_a is negative (silicon anode) and remains nearly unchanged within a definite region even if the surface is illuminated. s can be observed in this region only, because high current disturbs the measurements. Neither the conditions which determine the width of this region nor the current limiting process have been investigated so far.

In nearly all cases the correlation between s and V_a has been found to be V-shaped as shown in Fig. 4. The assumption suggests itself that these V-curves correspond somehow to the "Stevenson–Keyes curve", i.e. the bell-shaped dependence of s on the surface potential, ϕ_s as calculated by STEVENSON and KEYES[4]. They assume recombination centres of the Shockley–Read type to be present at the surface. If this supposition is correct, the applied voltage V_a must influence the surface potential ϕ_s by some sort of "field effect", an obvious assumption recently pointed out by DEWALD[1] and earlier by ENGELL and BOHNENKAMP[5] who measured the capacitance of germanium–electrolyte contacts.

THE SURFACE PHOTO EFFECT

In the present case, the assumption mentioned in the last paragraph has been verified by observing another effect greatly influenced by the surface potential ϕ_s: namely the surface photo effect. This effect generates a voltage between silicon and electrolyte due to the fact that equal numbers of excess electrons and holes generated in the bulk by light cause non-equal distribution of both carriers in the space charge layer at the surface as well as in the surface states. Therefore, a rearrangement of charges due to a change of the potential distribution is needed to maintain neutrality. In rough approximation—no surface states—it may be said that the surface photo voltage is caused by a charge separation of excess carriers in the field of the surface space charge layer. Thus high photo voltage should be expected if an inversion layer is present, and low voltage with opposite sign should be expected in the case of an accumulation layer.[6]

As shown in Fig. 6 the observation agrees qualitatively with this model: a negative value of V_a, i.e. silicon anode, makes the surface more p-type causing an inversion layer—high photo voltage—on n-type silicon and an accumulation layer—low

photo voltage—on p-type silicon. As expected, both voltages change their signs at different values of V_a. In the experimental arrangement (Fig. 2) the surface photo voltage is measured at the platinum wire. Using modulated light the photo voltage alternates and it can be separated therefore from the steady voltage V_a easily. If the light intensity is too high, the surface photo voltage becomes a non-linear function of the light intensity. Thus, the light intensity was adjusted and kept low for each of the measurements in Fig. 6.

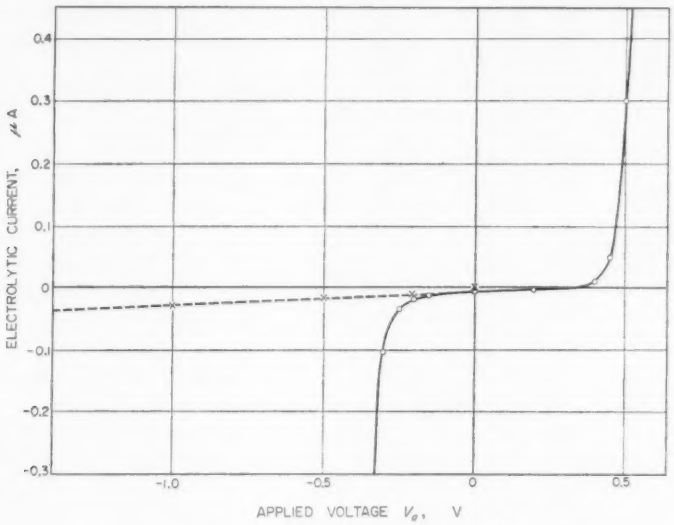


FIG. 5. Electrolytic current vs. V_a.

The surface photo voltage creates an alternating current in the electrolytic circuit causing a voltage drop across the base resistance of the silicon wafer. That means that an alternating voltage in phase with the photo voltage at the junction appears not only between the base contact of the wafer and the surface but also between the base contact and the junction. This voltage adds to the barrier photo effect and sometimes exceeds it. A disturbance of the recombination measurements, however, can be avoided by forming some sort of Wheatstone bridge with the two additional resistances R_1 and R_2 as shown in Fig. 7. If the ratio R_1/R_2 equals the ratio of barrier- and base-resistance the bridge is balanced and the wave analyser acting as the zero-indicator is not influenced by the bridge current, i.e. the surface photo current. On the other hand, the wave analyser remains the indicating instrument of the barrier photo voltage without losing sensitivity if R_1 is high enough. (In all cases R_1 was 1 MΩ and R_2 about 1 kΩ depending on the actual silicon wafer).

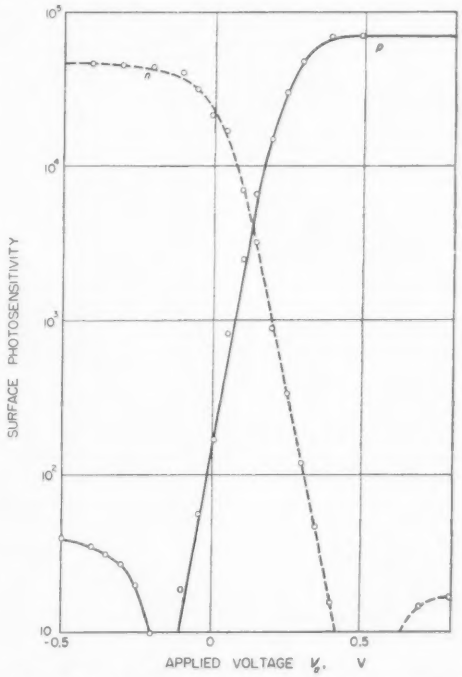


FIG. 6. Surface photo-voltaic sensitivity effect vs. V_a.

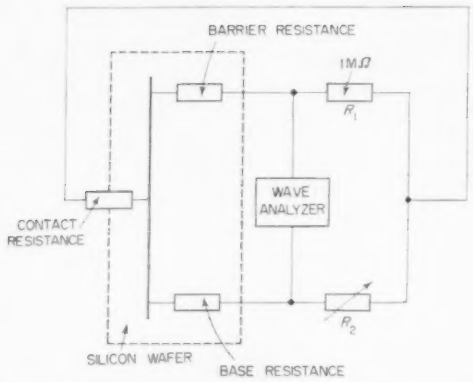


FIG. 7. Scheme of the Wheatstone bridge-like connection avoiding the disturbance of the barrier photo effect due to the surface photo current.

THE INFLUENCE OF AN OXIDE LAYER

From the observations mentioned above it follows that the applied voltage V_a affects the surface potential ϕ_s and that the covered range of ϕ_s is principally through which the potential drop occurs:

(1) the space charge region within the silicon,
(2) the oxide layer at the surface,
(3) the so-called "Helmholtz layer" caused by the finite diameter of the ions in the electrolyte,
(4) the space charge layer within the electrolyte (Gouy layer).

The thickness of each of the first two layers may lie in the order of microns. Therefore the voltage drop across the Helmholtz layer (thickness in the order of Å) may be neglected, as may the drop across the Gouy layer if the electrolyte is not too diluted. In a very rough approximation the ratio of the voltages across the space charge layer and across the oxide layer may equal the ratio of the thicknesses of both. Then a definite variation of V_a causes a variation in ϕ_s which is larger the thinner is the oxide layer relative to the space charge layer. In a second rough approximation it

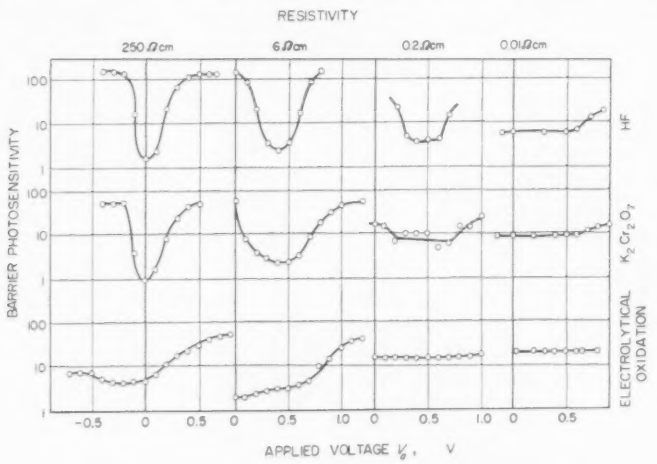


FIG. 8. Barrier photo effect vs. V_a with different silicon wafers after different surface treatments. From the left to the right decreasing resistivity (decreasing thickness of surface barrier layer), from top to bottom increasing thickness of the oxide layer.

larger than with field effect experiments at dry surfaces. The exact correlation between V_a and ϕ_s, however, is not known. At the silicon–electrolyte interface four different layers can be distinguished may be assumed that the width of the Stevenson–Keyes curve, related to ϕ_s, does not depend on the bulk resistivity of the silicon and that the width is not influenced by a surface oxidation—in other

words, that the type, not the number, of the actual recombination centres at the surface is almost the same. With these assumptions it follows that if the space charge layer is thick (high resistivity silicon) and the oxide layer is thin (etched surface), a relatively small variation of V_a is sufficient to cover the Stevenson–Keyes curve completely. Therefore the measured V-curve, related to V_a, should be narrow. The curve should be widened if the thickness of the space charge layer is decreased (low resistivity silicon) or if the oxide layer has grown thicker.

Corresponding measurements are shown in Fig. 8. Four wafers of silicon with different bulk resistivities have been investigated after three different surface treatments, namely

(1) etching with HF (thin oxide layer),
(2) oxidation with aqueous solution of $K_2Cr_2O_7$ (medium oxide layer),
(3) electrolytical oxidation with CH_3COOH (thick oxide layer).

The result corresponds to the expectations and therefore may justify the somewhat venturous assumptions.

The exact correlation between V_a and ϕ_s is not yet known. For this purpose measurements of the surface conductance and the contact capacitance have been started and preliminary results led to the assumption that, at least in some cases and in a limited region, the surface potential ϕ_s follows nearly exactly the applied voltage V_a.

Acknowledgements—This work was started during my association with the VALVO-GmbH-laboratory, Hamburg. To the head of this laboratory, Dr. H. TE GUDE, I am indebted for his kind support.

For valuable discussions I should like to thank Prof. D. POLDER, Dr. B. H. SCHULTZ, Dr. M. J. SPARNAAIJ, Dr. F. H. STIELTJES of the Philips Research Laboratories, Eindhoven, as well as several other colleagues. Also I wish to thank Miss I. MOSER for carrying out the experimental work very carefully.

REFERENCES

1. DEWALD J. F., *Semiconductor Chemistry* (edited by HANNAY N. B.). Reinhold, New York (1959).
2. HARTEN H. U. and SCHULTZ W., *Z. Phys.* **141,** 319 (1955).
3. HARTEN H. U., *Philips Res. Rep.* **14,** 346 (1959).
4. STEVENSON D. T. and KEYES R. J., *Physica, s'Grav.* **20,** 1041 (1954).
5. ENGELL H. J. and BOHNENKAMP K., *Z. Elektrochem.* **61,** 1184 (1957).
6. GARRETT C. G. B. and BRATTAIN W. H., *Phys. Rev.* **99,** 376 (1955).

VOL. 14 1960

VOL
14
196

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 227–232. Printed in Great Britain.

VOL. 14 1960

SECTION VI. THEORY

SURFACE STATES ASSOCIATED WITH ADSORBED ATOMS

T. B. GRIMLEY

Department of Inorganic and Physical Chemistry, University of Liverpool, Liverpool, England

Abstract—When a foreign atom is in interaction with a crystal surface, discrete electron levels may exist. These levels lie outside the normal crystal bands (including any surface bands) and the corresponding wave functions are large on the foreign atom and on the crystal atoms near it in the surface but fall to zero at points in the crystal remote from the foreign atom. Such levels allow for the formation of localized bonds between the foreign atom and the crystal. For semiconducting crystals they can provide surface traps for either electrons or positive holes. When the surface is completely covered by foreign atoms, bands of surface states may exist although no surface states exist when the surface is clean and no discrete levels are formed when only one foreign atom is present.

1. INTRODUCTION

THERE are two important problems in electronic structure to be solved before we can get a clear understanding of the surface properties of solids. First, there is the problem of a solid with a free surface, and second, there is that of the same solid with its surface contaminated by adsorbed atoms or molecules. The first of these has a history going back to TAMM's[1] paper in 1932, but the second has only recently received attention,[2, 3] and there is still a good deal to be discovered about it even with the simplest wave mechanical methods. If we use the simple LCAO (linear combination of atomic orbitals) approximation with a model crystal whose important electronic structure when there are no free surfaces consists of one band, derived in the LCAO approximation from a single atomic state, we can handle many problems involving chemisorbed atoms. For example, we can show how localized bonds are formed between a foreign atom and a crystal surface,[2, 3] and by treating the case where two foreign atoms are present, we can see that there are important interactions between chemisorbed atoms which actually take place through the crystal. KOUTECKÝ[2] has discussed the situation for two foreign atoms using a one-dimensional model for the crystal. However, there is no difficulty with a three-dimensional crystal, and the results will be mentioned in the present paper. In view of the important interaction effects exhibited when two foreign atoms are present, it is natural to consider the extreme case of a crystal surface completely covered by foreign atoms. This is actually a simpler problem than either of those mentioned above, and we shall discuss it in Section 5. It is obviously important to have a clear idea of what the terms "surface state" and "localized state" mean in our model. We begin therefore by examining briefly the electronic structure of a crystal with a free surface,[4] and that of a crystal with one foreign atom adsorbed.[3]

2. A CRYSTAL WITH A FREE SURFACE

Consider a simple cubic crystal in which the positions of the atoms are specified by vectors $\boldsymbol{m} = m_1\boldsymbol{\tau}_1 + m_2\boldsymbol{\tau}_2 + m_3\boldsymbol{\tau}_3$. m_1, m_2 and m_3 are integers, and $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$ and $\boldsymbol{\tau}_3$ are the vectors which map out the unit cell. Associated with each vector $\boldsymbol{m}$ we introduce an atomic orbital $\phi(\boldsymbol{r},\boldsymbol{m})$, and assume that any wave function $\psi(\boldsymbol{r})$ for the crystal can be expressed as

$$\psi(\boldsymbol{r}) = \sum_{\boldsymbol{m}} \phi(\boldsymbol{r},\boldsymbol{m})c(\boldsymbol{m}). \qquad (1)$$

Substituting this into the Schrödinger equation

$\mathscr{H}\psi = E\psi$, neglecting overlap between the atomic orbitals, and retaining only interactions between nearest neighbours, we arrive at the usual set of linear equations

$$(E-\alpha)c(\boldsymbol{m}) = \beta[c(\boldsymbol{m}+\boldsymbol{\tau}_1)+c(\boldsymbol{m}-\boldsymbol{\tau}_1)+ \\ +c(\boldsymbol{m}+\boldsymbol{\tau}_2)+c(\boldsymbol{m}-\boldsymbol{\tau}_2)+c(\boldsymbol{m}+\boldsymbol{\tau}_3)+c(\boldsymbol{m}-\boldsymbol{\tau}_3)]. \quad (2)$$

α is the Coulomb integral and β the resonance integral between nearest neighbours in the lattice. In the usual crystal problem we imagine that we have to deal with an infinite crystal which, however, is divided up into finite cubes each containing a large number of atoms. We then impose the condition (a boundary condition on the difference equation (2)) that the wave functions are the same at equivalent points in every cube. If there are N atoms along each cube edge, this condition is expressed by saying that

$$c(\boldsymbol{m}) = c(\boldsymbol{m} \pm N\boldsymbol{\tau}), \quad (3)$$

where $\boldsymbol{\tau}$ stands for either $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$ or $\boldsymbol{\tau}_3$. In this way, everything connected with the possible existence of free surfaces is removed from the problem. The solution of equations (2) and (3) is

$$c(\boldsymbol{m}) = \exp[i(m_1\theta_1+m_2\theta_2+m_3\theta_3)] \quad (4)$$

with θ_1, θ_2 and θ_3 integral multiples of $2\pi/N$, i.e.

$$\theta_1 = 2k_1\pi/N, \quad (-\tfrac{1}{2}N < k_1 \leqslant \tfrac{1}{2}N), \quad \theta_2 = 2k_2\pi/N, \\ (-\tfrac{1}{2}N < k_2 \leqslant \tfrac{1}{2}N), \theta_3 = 2k_3\pi/N, (-\tfrac{1}{2}N < k_3 \leqslant \tfrac{1}{2}N)$$

The corresponding energy levels are given by

$$E = \alpha+2\beta(\cos\theta_1+\cos\theta_2+\cos\theta_3). \quad (5)$$

We note that the wave functions (4) have periodic properties in all directions in the crystal.

For a finite crystal we obviously cannot impose the periodicity condition of equation (3). Suppose for example that we have a crystal confined by $\langle 100\rangle$ planes at $m_3 = 0$ and $m_3 = N$, but stretching to infinity in the other two directions. The vector $\boldsymbol{\tau}_3$ is perpendicular to the free surfaces; $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ are parallel to the free surfaces. In the directions of $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ we can still impose the periodicity condition (3) but we need new boundary conditions in the $\boldsymbol{\tau}_3$ direction. These new conditions must take account of the existence of the free surfaces. The simplest way to do this is to change the Coulomb integral on the surface atoms from its value α in the interior to a new value α'. Let $\boldsymbol{s}$ be a vector of the form $m_1\boldsymbol{\tau}_1+m_2\boldsymbol{\tau}_2$. Then $\boldsymbol{s}$ specifies the position of any atom in the surface $m_3 = 0$, and for these atoms

$$[E-\alpha']c(\boldsymbol{s}) = \beta[c(\boldsymbol{s}+\boldsymbol{\tau}_1)+c(\boldsymbol{s}-\boldsymbol{\tau}_1)+c(\boldsymbol{s}+\boldsymbol{\tau}_2) + \\ +c(\boldsymbol{s}-\boldsymbol{\tau}_2)+c(\boldsymbol{s}+\boldsymbol{\tau}_3)]. \quad (6)$$

This is the boundary condition at $m_3 = 0$. At the plane $m_3 = N$ we would also have a condition similar to equation (4). However, if N is large, this boundary condition cannot affect the conditions near $m_3 = 0$ in any important way, and we shall assume instead that the wave functions vanish on the plane $m_3 = N$. This defines our problem as that of a crystal with one free surface, namely that at $m_3 = 0$. The solution of equation (2) with our new boundary conditions is

$$c(\boldsymbol{m}) = \exp[i(m_1\theta_1+m_2\theta_2)]\sin(N-m_3)\theta_3. \quad (7)$$

θ_1 and θ_2 are defined in the same way as before, and the energy levels are still given by equation (5). The difference is that θ_3 is now one of the N roots of the equation

$$z+\cos\theta_3+\sin\theta_3\cot N\theta_3 = 0 \quad (8)$$

with

$$z = (\alpha-\alpha')/\beta. \quad (9)$$

For each root of equation (8) there are N^2 states corresponding to values of θ_1 and θ_2 in the range between $-\pi$ and π. Equation (8) has at least $N-1$ real roots, and according to equation (7) the corresponding wave functions are periodic in the crystal. If we write $\epsilon = (E-\alpha)/2\beta$, these $N^2(N-1)$ states lie in the range

$$-3 < \epsilon < 3 \quad (10)$$

which defines the familiar band of crystal states. The remaining root of equation (8) may also be real, and in this case the remaining N^2 states are also ordinary crystal states lying in the band (10). But if $|z| > 1+N^{-1}$, the remaining root of equation (8) has θ_3 of the form $i\xi$ or $\pi+i\xi$ with ξ real and positive. The wave functions for these states are not periodic in the $\boldsymbol{\tau}_3$ direction (perpendicular to the surface). Instead, the amplitude is large at the surface but falls steadily in the interior of the crystal. The N^2 levels of this sort constitute a *band of surface states*. The quantity ξ determines the spread of the wave functions into the interior of the crystal; the higher the value of ξ,

the more accurately are the surface states confined to the surface layer of atoms. A surface band with $\theta_3 = i\xi$ has $\epsilon - \cos\theta_1 - \cos\theta_2$ positive, and will be called a $\mathscr{P}$ band; a band with $\theta_3 = \pi + i\xi$ has $\epsilon - \cos\theta_1 - \cos\theta_2$ negative and will be called an $\mathscr{N}$ band. For large N, a $\mathscr{P}$ band exists if $z < -1$; for $z > 1$ there is an $\mathscr{N}$ band. In either case the ξ value is given by $|z| = \exp\xi$. If $\beta < 0$, a $\mathscr{P}$ band has its centre displaced below that of the normal crystal band whilst an $\mathscr{N}$ band has its centre above it. If $\beta > 0$, the situation is reversed. We note also that unless $|z|$ exceeds a certain critical value, the surface band overlaps the normal crystal band.

The surface states discussed above are what are usually called Tamm states. They arise because the energy parameters (in our case only the Coulomb integral) in the surface region are different from those in the interior. We know from the work of SHOCKLEY[5] and ARTMANN[6] that had we introduced more than one atomic orbital on each crystal atom so as to produce an overlapping band system then, even without modifying the energy parameters in the surface region, bands of surface states can be formed. These Shockley states are difficult to treat even in the simple LCAO approximation, and we shall not discuss them further.

VOL. 14 1960

3. A CRYSTAL WITH ONE ADSORBED ATOM

With the same notation as in Section 2 suppose there is a hydrogen-like foreign atom adsorbed over the surface crystal atom specified by $m_1 = m_2 = m_3 = 0$. The position of the foreign atom is specified by the vector $\boldsymbol{\lambda} = \lambda\boldsymbol{\tau}_3$, where λ is negative and non-integral, and associated with $\boldsymbol{\lambda}$ there is an atomic orbital $\phi(\mathbf{r},\boldsymbol{\lambda})$. We now have a similar problem to that of Section 2 except that the orbital $\phi(\mathbf{r},\boldsymbol{\lambda})$ must be included in summation in equation (1), and we need some new boundary conditions to take account of the presence of the foreign atom on the surface. Let α'' be the Coulomb integral on the foreign atom and let β' be the resonance integral between this atom and the surface crystal atom over which it is adsorbed. Then

$$[E-\alpha'']c(\boldsymbol{\lambda}) = \beta' c(\mathbf{0}). \tag{11}$$

The presence of the foreign atom will change the Coulomb integral on the surface atom over which it is adsorbed from its normal value α' to a new value $\alpha' + V$, but in the nearest neighbour approximation this is the only change to be made. Hence, our previous boundary condition (6) holds for all surface atoms except the one at the origin ($\boldsymbol{s} = \mathbf{0}$), and for this atom we have instead

$$[E-\alpha'-V]c(\mathbf{0}) = \beta' c(\boldsymbol{\lambda}) + \beta[c(\boldsymbol{\tau}_1) + c(-\boldsymbol{\tau}_1) + c(\boldsymbol{\tau}_2) + c(-\boldsymbol{\tau}_2) + c(\boldsymbol{\tau}_3)]. \tag{12}$$

With these new boundary conditions, the difference equation (2) does not have a simple solution. A solution (a Green's function solution) can however be obtained by the methods developed by BALDOCK[7] and by KOSTER and SLATER[8]. The results depend upon the parameter z which governs the occurrence of surface states for the crystal alone, and on three new parameters which specify the interaction between the foreign atom and the crystal. These parameters are

$$f = -V/\beta, \quad z' = (\alpha-\alpha'')/\beta, \quad \eta = \beta'/\beta. \tag{13}$$

Suppose first that $|z| < 1$ so that the crystal has no surface states. Then it turns out that the system foreign atom plus crystal may have *localized states*. The wave functions for these states are large either on the foreign atom, or on the crystal atom over which it is adsorbed, or on both, but fall to zero at points in the crystal remote from the foreign atom. The wave functions are not periodic in any direction in the crystal, and the corresponding energy levels lie outside the normal band of crystal states defined by equation (10). At the most two localized states can be formed, and depending on the values of the interaction parameters (13), there may be only one or even none. In the latter case, all the N^3+1 states of the system are non-localized, their wave functions are periodic in the crystal, and their energies lie in the band (10). Localized states will be called $\mathscr{P}$ or $\mathscr{N}$ states according as they have positive or negative values of ϵ. If $\beta < 0$, $\mathscr{P}$ states have energies lying below the normal crystal band, $\mathscr{N}$ states lie above it.

If $|z| > 1$, the crystal has a band of surface states, and this has an important effect on the positions of any localized states. Namely, any such states must lie outside the surface band as well as outside the main band of crystal states.

In concluding this brief survey, we note first that localized states allow for the formation of localized bonds between the foreign atom and the crystal, and second that with semiconducting

crystals, they can provide localized surface traps for either electrons or positive holes.

4. A CRYSTAL WITH TWO ADSORBED ATOMS

When two foreign atoms are present on the crystal surface, the wave functions for the system are either even or odd under the operation of inversion in the mid-point of the line joining the two atoms. In particular this applies to any localized states. When the two foreign atoms are infinitely far apart, any localized levels are doubly degenerate, but at any finite separation each level splits into an even and an odd level. If no surface states exist for the crystal alone, the even localized states lie further from the normal crystal band and the odd states lie closer to it than the original doubly degenerate level. The magnitude of the splitting increases as the distance between the two atoms is reduced, and depending on the position of the doubly degenerate level when the two atoms are far apart, there may be a critical value of the separation below which the odd localized state merges into the normal crystal band. When this happens the state is no longer localized. There is therefore the possibility that, although a localized state exists when one foreign atom is adsorbed, we do not get two localized states when two foreign atoms are present. This shows how the nature of the bond between one foreign atom and the crystal is modified by the presence of another atom adsorbed near it. It is important to realize that this effect is not due primarily to any direct interaction between the adsorbed atoms. The effect persists even when the resonance integral between the two adsorbed atoms is put equal to zero. There is an interaction taking place *through* the crystal by a certain delocalization of the bonding electrons in the crystal surface.

In view of these results we might speculate a little on the situation when the crystal surface is completely covered with adsorbed atoms. First we expect to find bands of surface states instead of the discrete localized states of a single adsorbed atom. Second, in view of the results when two foreign atoms are adsorbed, we would not expect that the surface bands would always contain as many levels as there are surface atoms. These speculations are in fact sound, but because surface states do not necessarily lie outside the normal crystal band whereas localized states always do, the completely covered crystal shows some new features. We shall therefore examine it in more detail.

5. A CRYSTAL SURFACE COMPLETELY COVERED WITH ADSORBED ATOMS

A simple cubic crystal has every atom in a $\langle 100 \rangle$ surface covered by a hydrogen-like foreign atom. We handle this problem in the simple LCAO approximation by changing the Coulomb integral from its value α in the interior of the crystal to α' on the crystal atoms at the surface, and to α'' on the atoms in the adsorbed layer. Also the resonance integral is β between nearest neighbours in the interior of the crystal, β' between a surface crystal atom and the adsorbed atom over it, and β'' between nearest neighbours in the adsorbed layer. These modifications of the energy parameters give two boundary conditions on the difference equation (2). The solution is simple. We have

$$c(\boldsymbol{m}) = \exp[i(m_1\theta_1+m_2\theta_2)]\sin(N-m_3)\theta_3 \quad (14)$$

for all atoms $\boldsymbol{m}$ in the crystal, and

$$c(\boldsymbol{\lambda}) = C\exp[i(m_1\theta_1+m_2\theta_2)] \quad (15)$$

for all atoms $\boldsymbol{\lambda}$ in the adsorbed layer. θ_1 and θ_2 are integral multiples of $2\pi/N$ just as in Section 2, but θ_3 is one of the roots of the equation

$$[z+\cos\theta_3+\sin\theta_3\cot N\theta_3][z'+2(1-\eta')\times \times(\cos\theta_1+\cos\theta_2)+2\cos\theta_3] = \eta^2. \quad (16)$$

Here

$$z = (\alpha-\alpha')/\beta,\ z' = (\alpha-\alpha'')/\beta,\ \eta = \beta'/\beta,\ \eta' = \beta''/\beta. \quad (17)$$

The energy levels are given by

$$\epsilon = \cos\theta_1+\cos\theta_2+\cos\theta_3$$

with $\epsilon = (E-\alpha)/2\beta$. We note that a crystal with a free surface is really a special case of the present problem. It is obtained when $\eta = 0$ by setting the first factor in equation (16) equal to zero.

Real roots θ_3 of equation (16) give energy levels lying in the normal crystal band ($-3<\epsilon<3$), and according to equation (14) the corresponding wave functions are periodic in all directions in the crystal. These are non-localized states associated with the crystal and the adsorbed layer. On the other hand, equation (16) may have roots with θ_3

of the form $i\xi$ or $\pi+i\xi$ with ξ real and positive. The corresponding wave functions fall steadily as we move into the crystal away from the surface. These are surface states associated with the adsorbed layer and the crystal surface. States with $\theta_3 = i\xi$ will be called $\mathscr{P}$ states, those with $\theta_3 = \pi+i\xi$ will be called $\mathscr{N}$ states, and in general we would expect to have bands of such states. A complete surface band will contain as many levels (N^2) as there are atoms in the crystal surface. Now the condition that, for example, a complete $\mathscr{P}$ band exists is that equation (16) has a root θ_3 of the form $i\xi$ for all values of θ_1 and θ_2. It is easy to see that this is not always possible, and that $\mathscr{P}$ states might exist only for a certain range of values of $(\cos\theta_1+\cos\theta_2)$. In this case only an *incomplete* $\mathscr{P}$ band is formed. Similar remarks apply to the $\mathscr{N}$ band. As an example of this sort of situation we illustrate in Fig. 1(a) the electronic structure at complete

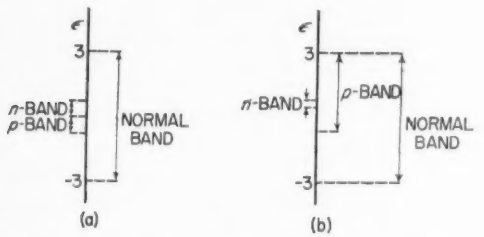


FIG. 1. Examples of the electronic structure at complete coverage.

(a) $z = z' = 0,\ \eta' = \frac{1}{2},\ \eta^2 = 1$
(b) $z = -1,\ z' = 0,\ \eta' = \frac{1}{2},\ \eta^2 = 1.$

surface coverage for $z = z' = 0,\ \eta' = \frac{1}{2}$ and $\eta^2 = 1$. There is an incomplete $\mathscr{P}$ band and an incomplete $\mathscr{N}$ band, each of width $1{\cdot}5|\beta|$, and these two surface bands are just touching at the centre of the normal crystal band. Of the extra N^2 levels which the adsorbed atoms bring to the system, those not in the surface bands are of course to be found as non-localized levels in the normal crystal band. A rather different situation is shown in Fig. 1(b). Here the interaction parameters are $z = -1$, $z' = 0, \eta' = \frac{1}{2}$ and $\eta^2 = 1$. There is an incomplete $\mathscr{N}$ band lying wholly within a complete $\mathscr{P}$ band. These are just two examples (among many) which confirm our speculation at the end of Section 4 that surface bands for a completely covered surface need not necessarily contain N^2 levels. Again we note that this result does not depend on any direct interaction between the atoms in the adsorbed layer. It persists even if $\eta' = 0$ so that the important interaction is actually taking place through the crystal. There is however, a more remarkable feature present in the examples quoted above. Namely, that we find bands of surface states at complete coverage with values of the interaction parameters for which no localized states exist when only a single atom is adsorbed. Again, this result is a rather general one, which persists even when $\eta' = 0$. It is another indication that the nature of the bonding between the adsorbed atoms and the crystal must change in going from a sparsely covered to a completely covered surface.

At the most two surface bands can exist, and these can be any combination of $\mathscr{P}$ and $\mathscr{N}$ bands. There may be only one band but provided that η' is not too near unity, a surface band of some sort (possibly incomplete) exists for all values of the interaction parameters. If η' is exactly unity, the situation is simple because the second term in the second bracket in equation (16) vanishes. This means that if a surface band exists at all, it is a *complete* band, and every state in it is characterized by the same value of ξ. There may be two surface bands or only one, and if $\eta^2 < 2$, there is a region round $z = z' = 0$ where no surface bands are formed. In this case a complete layer of adsorbed atoms is bonded to the crystal without any localization of the bonding electrons.

6. CONCLUSIONS

An attempt has been made to show how the problem of the electronic structure of a crystal with either a clean or a contaminated surface is handled with the simplest wave mechanical method. KOUTECKÝ[2, 9] has shown how these problems can be formulated in a rather general way by using Wannier functions instead of ordinary atomic orbitals. When this is done, the important features present in the simple LCAO approximation are retained. It seems clear therefore that simple methods can provide a good indication of the sort of results to be expected in a more exact theory. This is of course already known to be the case in many other fields, for example in dealing with the electronic structure of molecules.

The most pressing need at the moment is surely to extend our calculations to crystals with several

bands derived, in the LCAO approximation, from several atomic states. This does not present any formal difficulty, but the resulting equations are quite complicated, and it is not easy to extract any useful information from them. Another important problem is that of crystals with the diamond lattice. This lattice can be regarded as two interpenetrating face centred cubic lattices, and because of this, even a "one band model" shows an incipient band splitting in the sense that the normal crystal band has zero density of states at its centre. We also know from BALDOCK's[7] work that, if the crystal has a free surface, then even if the energy parameters for the surface atoms are not changed, the model shows a band of surface states of zero width with energy lying at the centre of the normal crystal band. In view of this result it would be very interesting to try to treat both clean and contaminated surfaces of such a crystal using a "two band model" with s and p atomic orbitals on the crystal atoms.

Returning now to the one band model, we have seen how localized states are formed when one foreign atom is adsorbed. These states are associated with both the foreign atom and the surface crystal atom over which it is absorbed, and we need scarcely add that they may be anionic (electron concentrated more on the adsorbed atom), cationic (electron concentrated more on the crystal) or even homopolar (electron shared equally between the adsorbed atom and the crystal). When two adsorbed atoms are present, the situation as regards localized states is rather similar to that with a homonuclear diatomic molecule except that there is an important interaction taking place through the crystal, and we have the requirement that localized states for the two adsorbed atoms must lie outside the normal crystal band. When the crystal surface is completely covered with adsorbed atoms, we find bands of surface states instead of discrete localized states. We have also seen that such surface bands (perhaps incomplete) may exist even though no discrete states are formed when only a single adsorbed atom is present. The reverse situation may also arise. Namely, that although there is a localized state when a single atom is adsorbed, we do not get a *complete* surface band when the surface is covered. This possibility stimulates an interesting speculation. Suppose that important binding of the adsorbed atom to the crystal surface must involve the formation of a localized state. Then to bind a complete layer we would need a complete band of surface states. In view of what has been said above, this may be impossible, and it appears therefore that there is a mechanism here for the formation of a saturated surface layer when only a fraction of the surface is covered.

REFERENCES

1. TAMM I., *Phys. Z. Sowjet.* **1,** 733 (1932).
2. KOUTECKÝ J., *Trans. Faraday Soc.* **54,** 1038 (1958).
3. GRIMLEY T. B., *Proc. Phys. Soc. Lond.* B**72,** 103 (1958).
4. GOODWIN E .T., *Proc. Camb. Phil. Soc.* **35,** 221 (1939).
5. SHOCKLEY W., *Phys. Rev.* **56,** 317 (1939).
6. ARTMANN K., *Z. Phys.* **131,** 224 (1952).
7. BALDOCK G. R., *Proc. Camb. Phil. Soc.* **48,** 457 (1952).
8. KOSTER G. F. and SLATER J. C., *Phys. Rev.* **95,** 1167 (1954).
9. KOUTECKÝ J., *Phys. Rev.* **108,** 13 (1957).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 233–240. Printed in Great Britain.

ON THE THEORY OF SURFACE STATES

J. KOUTECKÝ

Institute of Physical Chemistry, Czechoslovak Academy of Sciences, Prague

VOL. 14 1960

THE first condition for the development of a theory of the influence of surface states upon various physical and chemical processes taking place on the surface of solids consists in the achievement of clear and exact conceptions of the energy spectrum of an electron in a limited crystal. In this report we therefore concentrate our attention on just this fundamental problem of the theory of surface states, without considering the further important stages through which the theory must pass to be applicable to a quantitative elucidation of surface processes, such as the question of the statistical distribution of the electrons among the individual energy levels and the elaboration of a theory of the individual surface processes.

The theory of surface states has hitherto been treated only by methods based upon the one-electron approximation. An attempt has been made to treat the theory of chemisorption on metals by a method beyond this approximation. However, this is an isolated work, and deals only with a related problem.[1] For this reason our report also sets out from the one-electron approximation.

The long development of the theory of surface states from the first paper by TAMM[2], published in 1932, has supplied comparatively few works dealing with this subject, the fundamental significance of which is obvious. This has probably been caused by the circumstance that the difficult experimental technique of the investigation of surface processes has up to quite recent times not advanced far enough to present direct stimulation for the development of a theory in this field. The small number of publications[2–21] could not elucidate the pertinent problems, because it was found that the surface states of electrons in crystals are produced in various ways. The older works have therefore led to discrepancies, caused by unjustified generalizations of results depending on the calculating methods employed by the individual authors. Great progress in the classification of surface states was achieved by the work of SHOCKLEY[12]: However, some unclear points have not been eliminated even by newer publications setting out from a division of the surface states into Tamm states and Shockley states. We therefore believe that just at present—when the experimental technique of the investigation of pure surfaces has arrived at a real possibility of following surface states not caused by chemisorption or surface irregularities—it is fitting to present a survey of the fundamental conceptions of the states of electrons localized in a crystal surface. Special stress will here be laid upon a qualitative discussion, because just this is important for a correct elucidation of surface processes. The suitability of the approximate calculating methods employed, to which appropriate attention will be paid, also depends on the character of the surface states.

Of the Tamm surface states it is known that they are caused by the difference between the potential in the elementary cells lying directly under the crystal surface and the potential within the crystal. It is clear that the electron exhibits the tendency to remain in a region of an advantageous potential energy and, for this reason, an especially low energy will be encountered in states with an electron density materially higher at the surface of the crystal if the potential there is lower. If this deformation of the potential is sufficiently great, it can happen that in a certain state the electron is localized at the crystal surface in such a sense that the probability of finding the electron in this region is fixed, even if the dimensions of the crystal grow beyond any limits. We must bear in mind, however, that for non-localized volume-states exhibiting a higher energy the probability of the occurrence of an electron in the boundary region of the crystal

is lower than the probability of its occurrence within the crystal. For this reason the electron density at the crystal surface differs from the electron density within the crystal only when some of the energy levels under consideration remain unoccupied. If, on the other hand, there is a higher potential at the surface of the crystal the electrons in the states of lower energies tend to remain beyond this region so that the over-all electron density in volume states is higher within the crystal than under its surface. This difference is again compensated for by a surface state with higher energy. The formation of a surface state requires that the energy of such a state lie outside the band of allowed energies of the infinite crystal. Consequently, it is advantageous for the occurrence of Tamm surface states that there are considerable gaps between the individual bands of the volume states. It is therefore obvious that an especially suitable method for the investigation of Tamm surface states is the tight-binding method.

A characteristic feature of the theory of surface states is the existence of conditions for the occurrence of surface states that specify what difference there must be between the characteristics of the potential in the cells lying directly under the surface and of that in the cells within the crystal. Since the causes of the existence of Tamm surface states are in principle clear, the interesting problem is here only the dependence of these conditions on various factors. For our investigations we can select a model in which we consider a single band of energies, i.e. energies pertaining to wave functions for whose notation we need only the quantum number $\boldsymbol{k}$, characterizing the behaviour of the wave function under the operation of the elementary translation $\boldsymbol{\rho}$ of the crystal lattice:

$$\psi_k(\boldsymbol{r}-\boldsymbol{\rho}) = \exp[i(\boldsymbol{\rho},\boldsymbol{k})]\psi_k(\boldsymbol{r}). \qquad (1)$$

Such a model is suited only for crystals having a single atom in the elementary cell. But in this case we shall continue to speak of elementary cells in order not to lose the connection with implicit symmetrical properties and to indicate the possibility of generalization to crystals with several atoms in the elementary cell.

The wave function of an electron in an infinite crystal can be expressed as linear combination of the Wannier functions which are approximately localized in the regions of the individual elementary cells. We can imagine the finite crystal to be singled out of a cyclic crystal, for which the Born–Kármán conditions hold, by inserting a high potential in the part of the cyclic crystal which supplements the finite crystal to the cyclic one.[19] The wave functions of the electron in the finite crystal, which describe the state of this electron moving inside the finite crystal, can then be approximated with sufficient accuracy by the linear combination of the Wannier functions approximately localized in the elementary cells of the singled-out finite crystal and, perhaps, also of the Wannier functions pertaining to the elementary cells of the cyclic crystal at a short distance from the finite crystal. The formal similarity of the Wannier functions and the atomic orbitals of the tight-binding method shows that the tight-binding method can be considered to be a suitable approximation of the general expression of the wave function of the electron in the crystal. It is obvious that the form of the wave function, as well as that of the corresponding orbital energy, depends essentially on the form of the Wannier functions and on the dependence of the orbital energy on the quantum numbers k that characterize the behaviour of the wave function under elementary translations in the initial cyclic crystal. From this viewpoint the treatment of the problem of surface states, using the tight-binding method, appears to be a suitable first approximation: an approximation, however, whose consequences as well as the range of validity must be verified on more complicated models. For this reason, a more general theory was developed which, in addition to the mentioned formal similarity to the tight-binding method, affords the significant possibility of investigating more complicated models.[19, 22] This theory sets out methodically from the works of KOSTER and SLATER[23], BALDOCK[16] and LIFSHITZ[24]. From BALDOCK's procedure[16] it differs, apart from the use of Wannier functions, by the fact that it sets out from the orbital energies and the wave functions of an infinite (or a cyclic) crystal and not from the analogous quantities of a finite crystal not exhibiting a deformation of the potential in the cells lying under the surface.

Older papers, regardless of whether their authors had assumed a special potential shape[2, 5] or used the tight-binding method,[4, 8, 10, 15, 16] were restricted to the deformation of the potential inside

VOL. 14 1960

the elementary cells located directly under the surface. It was assumed that only the Coulomb integrals of the surface atom and, perhaps, the exchange integrals[4, 16] of these atoms, were different from the corresponding magnitudes for the atoms lying within the crystal. For this reason, it was logical to undertake in the framework of the tight-binding method a generalization for the simplest (viz. cubic) crystal, based upon the consideration of the deformation of the potential penetrating deeper below the surface.[19] By taking into account the dependence of the energy on the wave vector given by the relation

$$E_k = \epsilon_0 + 2\epsilon_1(\cos k_x a + \cos k_y a + \cos k_z a) \quad (2)$$

(a denotes the length of the elementary translation) it was possible to obtain a very simple equation for the calculation of the energy of the surface states. We give here the result for the case that the potential is still appreciably deformed in the cells adjoining the surface cells.

We denote the sought energy of the surface state by W and define

$$\epsilon_1 U_{00} = \int \tilde{a}_0 \Delta V a_0 \, dv, \quad \epsilon_1 U_{01} = \int \tilde{a}_0 \Delta V a_1 \, dv,$$
$$\epsilon_1 U_{11} = \int \tilde{a}_1 \Delta V a_1 \, dv, \quad \beta = W/\epsilon_1, \quad (3)$$

where a_0 denotes the Wannier function localized in the cell directly under the surface, a_1 the Wannier function in the second layer from the surface, and ΔV the deformation of the potential. The equation for calculating the energy reads:

$$\frac{\beta}{2}\left[1 + \sqrt{(1 - 4/\beta^2)}\right] = \frac{(U_{01}+1)^2}{\beta - U_{00}} + U_{11}. \quad (4)$$

By examination of this we easily find the conditions for the existence of the surface states, as given in Table 1. This table shows first of all that the surface states can exist even when the simple condition

$$|U_{00}| > 1 \quad (5)$$

does not hold, which must occur when the pertubation of the potential penetrates only into the surface cells, i.e. when $U_{01} = U_{11} = 0$ holds.

It is interesting that one of the surface states can exhibit a wave function with the maximum amplitude in the cells adjoining the surface cells. The energy of this subsurface state is lower than the energy of the states with the maximum amplitude on the surface if the potential in the surface cells is higher than inside the crystal.

A further possible generalization of the theory of the Tamm surface states consists in a more general assumption as to the form of the dependence of the energy on the wave vector. The physical interpretation of this generalization is afforded by considering the exchange integrals of the Wannier functions, or of the atomic orbitals pertaining to non-adjoining cells, or by the assumption of a more complicated geometric structure of the crystal. The simplest possible generalization of this kind is present by the following form of the dependence of the energy on the wave vector (21):

$$E_k = \epsilon_0 + 2\epsilon_1(\cos k_x a + \cos k_y a + \cos k_z a) +$$
$$+ 2\epsilon_2(\cos 2k_x a + \cos 2k_y a + \cos 2k_z a)_1 \quad (6)$$

if we consider the deformation of the potential in the first cell only (i.e. $U_{00} \neq 0$).

Instead of the single condition for the existence of Tamm surface states the following two conditions must be simultaneously satisfied:

$$|U_{00}| > |\epsilon_2/\epsilon_1|, \quad |U_{00} - \epsilon_2/\epsilon_1|^2 > |U_{00}|. \quad (7)$$

We see that the deformation of the potential must be greater than would correspond to the condition (5) when U_{00} and ϵ_2/ϵ_1 have equal signs, while the occurrence of surface states is impaired when these values exhibit opposite signs. For instance, if $|\epsilon_2/\epsilon_1| = 0{\cdot}25$ surface states are produced provided that $|U_{00}| > 1{\cdot}45$ and sign U_{00} = sign (ϵ_2/ϵ_1) or $|U_{00}| > 0{\cdot}25$ and sign $U_{00} = -\text{sign}\,(\epsilon_2/\epsilon_1)$. Besides this we notice that the coefficients in the wave function of the surface states are for some values of the parameters an oscillating function of the distance from the surface; but this periodic function is modulated by an exponentially descending function. Finally, if sign U_{00} = sign (ϵ_2/ϵ_1), the absolute value of the energy of the surface state, calculated for the energy given by equation (6), is greater than if we neglected the magnitude ϵ_2: On the other hand, if sign $U_{00} = -$sign (ϵ_2/ϵ_1), this value is smaller than the value calculated for the energy according to equation (2).

Still more complicated assumptions as to the energy of surface states were made by MAUE[6] and BALDOCK[16]. However, the discussion of the results cannot be very clear due to the complexity of these assumptions.

Table 1. *Examples of conditions for the existence of surface states when* $U_{11} = 0$ (*cf. eq.* 3). *According to Refs.* (19) *and* (35)

No.	Occurrence of surface states	Conditions
1	No surface states	$-6<U_{00}<6$ and $\lvert U_{01}+1\rvert^2<(U_{11}+1)(U_{00}+6)$ and $\lvert U_{01}+1\rvert^2<(U_{11}-1)(U_{00}-6)$
2	No surface states below the band of volume state energies	$U_{00}>-6$ and $\lvert U_{01}+1\rvert^2<(U_{11}+1)(U_{00}+6)$
3	Existence of a band of surface states below the band of volume state energies. Number of surface states is smaller than the number of surface cells.	$U_{00}>-6$ and $\lvert U_{01}+1\rvert^2>(U_{11}+1)(U_{00}+6)$
4	Existence of a band of surface states below the band of volume state energies. Number of surface states equals the number of surface cells.	$U_{00}\leqslant -6$ or $U_{00}>-6$ and $\lvert U_{01}+1\rvert^2>(U_{11}+9{\cdot}9)(U_{00}+6)$
5	Below the band of volume state energies lies a band of energies due to the surface states and a band of energies due to the subsurface states. In both cases the number of states equals the number of surface cells.	$U_{00}<-6$ and $\lvert U_{01}+1\rvert^2<(U_{11}+9{\cdot}9)(U_{00}+6)$

If the energy of the surface state is near the energies of the volume states falling into the next allowed band, it is necessary, however, even when considering the Tamm states, to take into account the Wannier functions pertaining to the energies of this band, when expressing the wave function of the surface state. Surface states may be absent for an especially great deformation of the potential of the surface for the reason that the energy of the surface state penetrates into the adjoining allowed band of volume states. An exception is presented by the case where the volume wave functions of neighbouring bands belong to such various irreducible representations of the symmetry group of the crystals that an interaction between the wave functions of both bands sets in by the influence of the perturbation surface potential which, of course, exhibits the symmetry of the crystal surface. The penetration of such a surface state into a fully occupied band can result in a charging of the surface under consideration.[25]

The hitherto mentioned methods neglect the interaction of the electrons even within the frame of the Hartree–Fock self-consistent method. However, in the first paper on surface states, TAMM[2] already pointed out the possible influence of the interaction of the electrons upon the occurrence of surface states. It is therefore natural to attempt to generalize the method of investigating surface states in such a way that the condition of self-consistency is satisfied. Concrete calculations can be carried out on the basis of a general theory[20] which, by analogy with the theory of electronic spectra of organic molecules, takes as eventually unoccupied states the solutions of the Hartree–Fock equations pertaining to higher energies than the energies of occupied states. Such calculations, using simplifications common in the theory of organic compounds, are just being carried out.[26]

From these considerations it is obvious that the various generalizations relate, of course, only to those quantitative criteria (for the existence of surface states) whose character does not significantly change if the subsurface states are not considered as a qualitatively new kind of surface state and if the possibility is not considered that a greater number of localized states may exist if the potential deformation extends deeper below the surface. Since the exchange integrals, Coulomb integrals and the integrals characterizing the perturbation

VOL. 14 1960

potential in the cells near the surface can evidently be magnitudes of the same order, these generalizations show that we cannot undertake a general consideration of the existence or probable non-existence of Tamm states—e.g. based upon a discussion of the relations, equation (5).

Many more nuclear points are encountered in the theory of the second group of surface states, which are termed Shockley states.[6, 12, 14, 15, 18] Only quite recently we have commenced to achieve clearer conceptions of the formation of these states, which occur only in the case of an interaction of the wave functions pertaining to two neighbouring bands of allowed energies of the electron in an infinite crystal. This interaction must be rather strong, which is the case for interatomic distances at which the wave functions of the two bands would be degenerate were it possible to neglect the interaction. However, this degeneracy is removed by the interaction: the energy of the bonding states is decreased and that of the anti-bonding states increased. If we plot the allowed energies of the infinite crystal against the interatomic distance, the region in which these conditions are satisfied is characterized by smaller interatomic distances than would correspond to an intersection of the bands of allowed energies (Fig. 1).[6, 18] For small distances between the atoms of which the crystal is built up it is known that the energy levels pertaining to the atomic orbitals spread into bands, and the mentioned conditions indicate that this must be so considerable as to present the possibility of overlapping. The removal of the degeneracy which thus could take place and the formation of a new gap between the bands of volume energies depend, however, on the special properties of the wave functions pertaining to both bands. But the question of necessary and sufficient conditions for the formation of the mentioned gap has hitherto not been investigated to a satisfactory degree. The most valuable studies in this respect are presented by the papers of SHOCKLEY[12] and LIPPMANN.[18] The second of these publications, however, deals only with a one-dimensional potential and shows that for such a potential overlapping of the bands can occur only then when the energy of the electron exceeds the highest value of the potential in the elementary cell of the crystal. This result is comprehensible, because the wave functions for the one-dimesional potential are characterized by a quantum number describing the behaviour of the wave function to the elementary translation and by a quantum number which is associated with the energy of the electron in one isolated elementary one-dimensional cell. No further quantum numbers exist in this case, and the individual band can be characterized by the latter quantum number. An overlapping of the bands can take place only then when the potential on the boundary of the individual elementary cells decreases to such an extent, that the condition derived by LIPPMANN is satisfied. From this it follows, concerning the selection of the calculating method, that for a one-dimensional potential the method of nearly free electrons is suitable for the study of the surface states (it was the procedure by which MAUE[6] first obtained Shockley states) while the tight-binding method is not permissible. Intersection of the bands is, of course, a necessary but not at all a sufficient condition, which also follows from the calculations using a special shape of the potential, equation (14). It is obvious that here the symmetry properties of the wave functions of both bands play an important part. Also SHOCKLEY[12] assumed that prior to intersection of the bands both the anti-symmetric part of the wave function for the bottom of the upper band and the symmetric part of the wave function for the top of the lower band have vanishing logarithmic derivatives at the edges of the elementary cell.

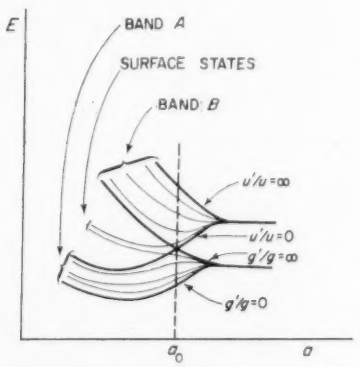


FIG. 1. Dependence of energies of a finite crystal model on the interatomic distance according to SHOCKLEY[12] and LIPPMANN[18]. g means the symmetrical and u the antisymmetrical part of the wave function.

If the potential is three-dimensional, i.e. if we must consider at least the potential in one elementary cell to be three-dimensional, regardless of whether these cells form a one-dimensional chain or a complete crystal, it is clear that the wave functions pertaining to different bands are characterized by further quantum numbers. The proximity of two bands, however, need not be caused by the circumstance that the potential on the boundary of two elementary cells would be lower than the energy of the considered electron. In this case it is possible to investigate Shockley states by means of the tight-binding method, as it was done by ARTMANN[15]. The use of this method has besides its simple formulation a considerable advantage in its clearness, which enables us to find the connections between the concepts of the theory of surface states and the concepts employed in surface chemistry. ARTMANN[15] investigated the model of a crystal in which a hybridization between the functions of s and p character must be assumed. This means that the wave functions of the electron in the isolated atoms of which the crystal is built up are symmetric atomical orbitals around the nucleus of the atom, or they are antisymmetric with respect to reflection in the plane parallel to the surface. If the energies pertaining to these atomic orbitals lie sufficiently near to one another and if the distance of these levels from the other energy levels is sufficiently great, it will be sufficient for the tight-binding approximation to consider only the hybridization between these atomic orbitals. It is obvious that this model, despite its rather unrealistic simplicity, is well suited for studying the influence of the symmetry of the initial wave functions upon the existence of Shockley surface states. ARTMANN[15], however, made assumptions which, even for this simple model, did not permit him a more general selection of the parameters entering the theory, and this circumstance has somewhat restricted the possibilities utilizing his work as a base for the interpretation of Shockley states. For this reason, a general method for the study of surface states was applied to ARTMANN's model with a more general selection of the parameters.[34]

Since the energy of the Shockley state lies in a narrow gap between two neighbouring bands of the energies of volume states, it is clear that the deformation of the potential in the proximity of the surface must not be too great or else the energy of the surface state will penetrate into one of these bands.[27] For this reason, the characteristic feature of Shockley states, in distinction to Tamm states, is just the circumstance that they can be brought about by a mere restriction of the crystal without a further deformation of the potential inside the crystal. Consequently, the investigation of the general character of the Shockley surface states can be restricted to the case of zero deformation of the potential. The results of this study are given in an accompanying paper.[34] For this reason, we present here only the most essential results.

If the values of the exchange integrals between the atomic orbitals are so great that intersection of the bands takes place, a considerable part of the volume state wave functions of the lower band are hybridized to such an extent that a high electron density is produced in those parts of the border cells which are farthest from the surface. This results in a maximum overlapping between neighbouring atoms which, according to the fundamental conceptions of the theory of the chemical bond, causes the diminution of the energy of these states.

In the states of the upper band, on the contrary, the hybridization causes a higher electron density in the outer parts of the surface cells. The wave functions of the surface states are then distinguished by a hybridization which effects a higher electron density in the outer parts of the surface cells (Fig. 2, cf. Ref. 28). These states, in which the electrons are localized on the crystal surface, under ordinary conditions are not fully occupied. From this consideration it follows that in the case of hybridization between atomic orbitals with

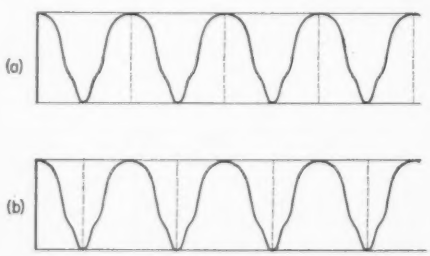


FIG. 2. Definition of the elementary crystal cell for an arbitrary one-dimensional shape of potential: (a) for interatomic distance $a > a_0$; (b) for $a < a_0$. The definition of the quantity a_0 is given on Fig. 1.

VOL. 14 1960

different additional quantum numbers the Shockley states are a physical description of the bonding states in the surface of the crystal, which in the chemistry of surfaces are termed free valences, i.e. unsaturated valences of the surface atoms. Thus we arrive at a clear interpretation of the Shockley surface states.

STATZ[14] and LIPPMANN[18] pointed out the connections between the Tamm and the Shockley surface states in the case of a one-dimensional potential. The limitation of the elementary cell is obviously to a certain extent arbitrary. In the region of interatomic distances shorter than the distance at which intersection of the bonds of volume states arises we must delimit the elementary cell in such a way that at the surface only a half of the border elementary cell is left (cf. Fig. 2). For this reason, we can formally conceive the difference between these half-cells at the surface and the complete cells within the crystal as a deformation necessary for the occurrence of Tamm surface states. From

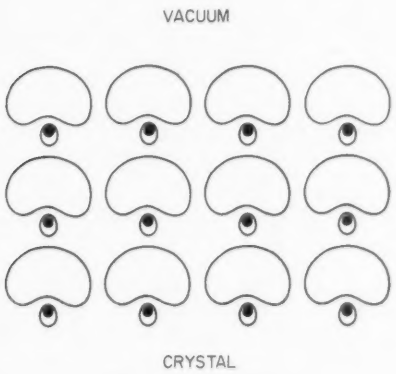


FIG. 3. Schematic representation of the hybridization of atomic s and p_z orbitals for the surface Shockley states (cf. Ref. 28).

this viewpoint it is evident, too, that in the case of Shockley surface states for a one-dimensional potential the number of states, whose energy falls into each of the volume bands between which the energy level of the Shockley state has been produced, must necessarily be reduced by one.[12, 18] If the number of electrons at disposal for occupying the energy levels equals the number of elementary cells, the surface states are occupied only to one half. This result, obtained for a one-dimensional potential, is probably not generally valid for a three-dimensional potential. According to STATZ[14], the cellular method provides that for his model the Shockley surface states originate from the upper band only. In the model suggested by GOODWIN and ARTMANN the upper and lower bands have one state less than the member of elementary cells[34]: Consequently, we may assume that one of the Shockley states originates from the upper and the other one from the lower band of energies (ARTMANN'S[15] statement that both of them originate from the upper band is erroneous).

From these facts it follows that Shockley states, caused by the proximity of volume bands characterized by quantum numbers of different kinds, can behave in various ways. If the elementary cell contains more than one atom, the divergence of the volume state bands can be caused by the different chemical nature of these atoms, or also by the circumstance that they are situated in non-equivalent positions in the elementary cell. It seems that here, too, the bond conditions specific for the surface can produce changes in the hybridization, which manifest themselves by the appearance of surface states whose creation requires no essential deformation of the potential of the surface. This opinion is supported by BALDOCK'S investigation[16] of the occurrence of surface states on more complicated two-dimensional crystal models.

The free valences of the crystal surface are of fundamental significance for the explanation of the nature of surface reactions and chemisorption. For instance, in WOLKENSTEIN'S interpretation of chemisorption on semiconductors it is assumed that the function of free valences in the surface is performed by electrons lifted into the conduction band.[29] The free valences created by special bond conditions in the surface, however, are not less important, as already pointed out in the literature (cf. e.g. Ref. 30). For a quantitative description of the qualitative concept of a free valence of this type expression in the terminology of the Shockley states may be advantageous. Here it will be adequate to point out the considerable formal similarity of the quantum-mechanical treatment of the problem of chemisorption and the problem of surface states, although we do not intend to deal in our report with the theory of chemisorption in any detail.[29, 31–33]

From the given principal features of the present stage of the theory of surface states it is evident that we have already achieved some results which can be used as a basis for the development of fundamental conceptions of the character of the energy spectrum of a finite crystal and of the properties of the wave functions in the proximity of a crystal surface. It seems desirable, however, to treat still more complicated crystal models which would more closely approximate real crystals and especially crystals interesting from the practical viewpoint. The development of a general mathematical method, together with the possibilities offered by automatic computers, renders calculations of this kind feasible. In addition to this, it will be necessary to analyse more generally the connection between the occurrence of the Shockley states and the symmetry properties of the involved Wannier functions for the infinite crystal. Another interesting question would be the investigation of different criteria for the occurrence of Shockley states which result from the hybridization of Wannier functions characterized by quantum numbers of different kinds. We believe that work in both of these directions would materially contribute to the qualitative, and, perhaps, the quantitative interpretation of the physical and chemical processes taking place on the surface of semiconductors and metals.

REFERENCES

1. Bonts-Bruevits V. L. and Glasko V. B., *Vest Mosk. Univ., ser. mat.* **91** (1958).
2. Tamm Ig., *Phys. Z. Sowjet.* **1,** 733 (1932); *Z. Phys.* **76,** 849 (1932).
3. Fowler R. H., *Proc. Roy. Soc.* A**141,** 56 (1933).
4. Ryzhanov S., *Z. Phys.* **89,** 806 (1934).
5. Sokolov A. A., *Z. Phys.* **90,** 520 (1934).
6. Maue A. W., *Z. Phys.* **94,** 717 (1935); *Naturwissenschaften* **38,** 648 (1934).
7. Sokolov A. A., *Zh. eksp. teor. fiz.* **6,** 807 (1936).
8. Center E. M., *Zh. eksp. teor. fiz.* **8,** 682 (1938).
9. Goodwin E. T., *Proc. Camb. Phil. Soc.* **35,** 205 (1939).
10. Goodwin E. T., *Proc. Camb. Phil. Soc.* **35,** 221 (1939).
11. Goodwin E. T., *Proc. Camb. Phil. Soc.* **35,** 232 (1939).
12. Shockley W., *Phys. Rev.* **56,** 317 (1939).
13. Coulson C. A. and Baldock G. R., *Disc. Faraday Soc.* **8,** 27 (1950).
14. Statz H., *Z. Natur.* **5a,** 534 (1950).
15. Artmann K., *Z. Phys.* **131,** 244 (1952).
16. Baldock G. R., *Proc. Camb. Phil. Soc.* **48,** 457 (1952).
17. Lifshitz I. M. and Pekar S. I., *Usp. fiz. nauk* **56,** 531 (1955).
18. Lippmann B. A., *Ann. Phys., Lpz.* **2,** 16 (1957).
19. Koutecký J., *Phys. Rev.* **108,** 13 (1957).
20. Koutecký J., *Czech. čas. fys.* **8,** 153 (1958); *Czech. J. Phys.* **8,** 148 (1958).
21. Tomášek M. and Koutecký J., *Czech. J. Phys.*, in press.
22. Koutecký J. and Fingerland A., *Dokl. Akad. Nauk. SSSR* **125,** 841 (1959).
23. Koster G. F. and Slater J. C., *Phys. Rev.* **95,** 1167 (1954); *Phys. Rev.* **96,** 1208 (1954).
24. Lifshitz I. M., *Zh. exp. teor. fiz.* **17,** 1017 (1947); *Zh. exp. teor. fiz.* **17,** 1076 (1947).
25. Handler P. and Portnoy W. M., *Electronic Surface States and Cleaned Germanium Surface*, preprint.
26. Koutecký J., *Czech. J. Phys.*, to be published.
27. Tomášek M., *Czech. J. Phys.*, to be published.
28. Moffitt W. E. and Coulson C. A., *Phil. Mag.* **38,** 634 (1947).
29. Wolkenstein F. F., *Advanc. Catalys.* **9,** 807 (1957).
30. Dowden D. A., *Chemisorption* p. 3. Butterworths, London (1957).
31. Koutecký J., *Z. Elektrochem.* **60,** 855 (1956).
32. Koutecký J., *Trans. Faraday Soc.* **54,** 1038 (1958).
33. Grimley T. B., *Proc. Phys. Soc. Lond.* **72,** 103 (1958).
34. Koutecký J. and Tomášek M., this Conference, following paper.
35. Koutecký J. and Fingerland A., *Coll. Trav. Chim. Tchécosl.* **25,** 1 (1960).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 241–247. Printed in Great Britain.

CONTRIBUTION TO THE THEORY OF THE SHOCKLEY SURFACE STATES—I

GENERAL FORMULATION AND THE CASE OF ZERO DEFORMATION OF THE POTENTIAL ON THE GOODWIN–ARTMANN MODEL

J. KOUTECKÝ and M. TOMÁŠEK

Institute of Physical Chemistry, Czechoslovak Academy of Sciences, Prague

VOL. 14 1960

Abstract—Within the framework of the tight-binding method, general equations have been set up for the calculation of orbital energies and wave functions of the electron in a crystal model in which an interaction between atomic s orbitals and p orbitals is assumed. Simple explicit expressions are derived for the case of zero deformation of the potential of the crystal surface. From a discussion of the hybridization of the wave functions of the volume and surface states the connection with the concept of free valence localized in the crystal surface follows.

1. INTRODUCTION

GOODWIN[1] and ARTMANN[2] investigated a crystal model in which it must be assumed that the wave function of the surface state is, in the tight-binding method, expressible as a linear combination of the atomic orbitals of s and p character. They considered only the p_z orbitals, taking as z axis the coordinate axis perpendicular to the crystal surface under investigation. GOODWIN[1] discussed only the existence of the Tamm states,[3] for he assumed that under conditions permitting the occurrence of Shockley states[4] the method of tight binding is not applicable. ARTMANN[2] noticed the fact that in the case of a three-dimensional potential the proximity of two bands need not be due to the circumstance that the energy of the electron would be higher than the level of the potential barrier between the atoms of the crystal (cf. Ref. 5). However, he restricted his investigation to a very special selection of the Coulombic and exchange integrals characterizing the problem. A more general treatment of this model, which would permit a more thorough-going discussion of the conditions of the existence and character of the Shockley surface states and their association with the properties of the initial atomic orbitals, is the objective of the present communication. Since the inclusion of the p_x and p_y orbitals would probably not lead to essentially new results from the viewpoint of a general discussion, they were for reasons of simplicity not taken into account, just as with the above-mentioned authors.[1, 2] We also neglected the exchange integrals between the atomic s orbitals and p_z orbitals for the atoms located in the same planes parallel to the surface, which due to symmetry, are smaller than the other exchange integrals.

2. ASSUMPTIONS AS TO THE WAVE FUNCTIONS AND ORBITAL ENERGIES OF THE ELECTRONS IN AN INFINITE CRYSTAL

Let us assume the wave function of the electron in an infinite cubic crystal in the form of the following linear function ($\boldsymbol{\rho}$, $\boldsymbol{\sigma}$, $\boldsymbol{\tau}$ = elementary translations of the crystal, α = I, II):

$$\psi^{\alpha}_{\xi\eta\zeta} = \frac{N^{-3/2}}{1+|\lambda^{\alpha}_{\xi}|^2}\sum_{l,m,n}\exp\left[i(l\xi+m\eta+n\zeta)\right]\times$$

$$\times[a_s(\boldsymbol{r}-l\boldsymbol{\rho}-m\boldsymbol{\sigma}-n\boldsymbol{\tau})+\lambda^{\alpha}_{\xi}a_p(\boldsymbol{r}-l\boldsymbol{\rho}-m\boldsymbol{\sigma}-n\boldsymbol{\tau})]. \quad (1)$$

The functions a_s, a_p are atomic orbitals, for which holds:

(i) The overlap integrals of these are throughout equal to 0.

(ii) If $\boldsymbol{H}$ is the Hamiltonian of the electron in an

infinite crystal and H_{lmn} the Hamiltonian that satisfies the characteristic equation ($\nu = s, p$):

$$(H_{lmn} - E_\nu) a_\nu(\mathbf{r} - l\boldsymbol{\rho} - m\boldsymbol{\sigma} - n\boldsymbol{\tau}) = 0 \qquad (2)$$

then, in compliance with our remarks in the introduction to this paper, we put

$$\int \bar{a}_s(\mathbf{r} - l'\boldsymbol{\rho} - m'\boldsymbol{\sigma} - n'\boldsymbol{\tau})[H - H_{lmn}] a_s(\mathbf{r} - l\boldsymbol{\rho} - m\boldsymbol{\sigma} - n\boldsymbol{\tau})\, dv$$

$$= -\alpha_s \delta_{ll'}\delta_{mm'}\delta_{nn'} - \gamma_s[(\delta_{l',l+1} + \delta_{l',l-1})\delta_{mm'}\delta_{nn'} + \delta_{ll'}(\delta_{m',m+1} + \delta_{m',m-1})\delta_{nn'} + \delta_{ll'}\delta_{mm'}(\delta_{n',n+1} + \delta_{n',n-1})],$$

$$\int \bar{a}_p(\mathbf{r} - l'\boldsymbol{\rho} - m'\boldsymbol{\sigma} - n'\boldsymbol{\tau})[H - H_{lmn}] a_p(\mathbf{r} - l\boldsymbol{\rho} - m\boldsymbol{\sigma} - n\boldsymbol{\tau})\, dv$$

$$= -\alpha_p \delta_{ll'}\delta_{mm'}\delta_{nn'} + \gamma_p(\delta_{l',l+1} + \delta_{l',l-1})\delta_{mm'}\delta_{nn'} - \gamma'_p\delta_{ll'}[(\delta_{m',m+1} + \delta_{m',m-1})\delta_{nn'} + \delta_{mm'}(\delta_{n',n+1} + \delta_{n',n-1})],$$

$$\int \bar{a}_s(\mathbf{r} - l'\boldsymbol{\rho} - m'\boldsymbol{\sigma} - n'\boldsymbol{\tau})[H - H_{lmn}] a_p(\mathbf{r} - l\boldsymbol{\rho} - m\boldsymbol{\sigma} - n\boldsymbol{\tau})\, dv = -\gamma_{sp}(\delta_{l',l+1} - \delta_{l',l-1})\delta_{mm'}\delta_{nn'},$$

where δ_{ij} is the Kronecker delta.

By the usual variation method we obtain the following relations for the coefficients λ_ξ^α:

$$X = -2\cos\xi + 2i\mu\lambda_\xi^\alpha \sin\xi,$$
$$\lambda_\xi^\alpha Y = -2i\nu \sin\xi + 2\lambda_\xi^\alpha \cos\xi, \qquad (4)$$

where

$$X = \frac{E - \eta_s}{\gamma_s} = \gamma_s^{-1}[E - E_s - \alpha_s - 2\gamma_s(\cos\eta + \cos\zeta)],\ \mu = \frac{\gamma_{sp}}{\gamma_s},$$
$$Y = \frac{E - \eta_p}{\gamma_p} = \gamma_p^{-1}[E - E_p - \alpha_p - 2\gamma'_p(\cos\eta + \cos\zeta)],\ \nu = \frac{\gamma_{sp}}{\gamma_p}, \qquad (5)$$

E denoting the energy of the electron in an infinite crystal.

By elimination of the magnitude λ_ξ^α from these equations we obtain the relation for the energy of the electron E in the form

$$(X + 2\cos\xi)(Y - 2\cos\xi) = 4\mu\nu \sin^2\xi. \qquad (6)$$

The solution of this quadratic equation leads to

$$E^{\mathrm{I,II}}(\xi) = E - \tfrac{1}{2}(\eta_s + \eta_p)$$
$$= (\lambda_p - \gamma_s)\cos\xi \pm 2[(q + q_0\cos\xi)^2 + \gamma_{sp}^2 \sin^2\xi]^{1/2} = E'(\xi) \pm E''(\xi) \qquad (7)$$

where we have put

$$q = \tfrac{1}{4}(\eta_p - \eta_s),\ q_0 = \tfrac{1}{2}(\gamma_s + \gamma_p). \qquad (8)$$

For the coefficients λ_ξ^α we finally obtain

$$\lambda_\xi^{\mathrm{I,II}} = i\gamma_{sp}^{-1}(\sin\xi)^{-1}[q + q_0\cos\xi \pm \tfrac{1}{2}E''(\xi)]. \qquad (9)$$

By differentiating the expression (7) for the energy of the electrons in the infinite crystal we find that it exhibits its extreme values for

$$\sin\xi = 0 \qquad (10)$$

eventually also for

$$\cos\xi = -\frac{q\, q_0}{q_0^2 - \gamma_{sp}^2} \pm \frac{\gamma_{sp}(\gamma_s - \gamma_p)}{2(q_0^2 - \gamma_{sp}^2)} \sqrt{\frac{q_0^2 - q^2 - \gamma_{sp}^2}{\gamma_s\gamma_p - \gamma_{sp}^2}} \qquad (11)$$

For the first case the energy values are given by the relation

$$E_{+-}^{\mathrm{I,II}} = \pm(\gamma_p - \gamma_s) \pm 2|q \pm q_0| \qquad (12)$$

For the second case the energy values are

$$E^{\mathrm{III,IV}} = \frac{1}{q_0^2-\gamma_{sp}^2}\,[qq_0(\gamma_s-\gamma_p)\pm \pm 2\gamma_{sp}(\gamma_s\gamma_p-\gamma_{sp}^2)^{1/2}(q_0^2-q^2-\gamma_{sp}^2)^{1/2}]. \quad (13)$$

The expression for the extreme values of the volume state energies given in the last equation has a sense for such values of q for which the right side of equation (11) has an absolute value less than unity and for which the square root in equation (13) is at the same time real. By a simple analysis of these conditions we find that the limits of the bands for $q > q_0$ are given by the relations (11).

For $q \leqq q_0$ and $\gamma_s\gamma_p > \gamma^2$, the lower limit of the upper band is in the interval Δ_{1q}:

$$-q_0+\frac{\gamma_{sp}^2}{\gamma_p} \leqslant q \leqslant q_0-\frac{\gamma_{sp}^2}{\gamma_s} \quad (14)$$

defined by the relation (13) with the upper sign; the upper limit of the lower band is given by the relation (13) with the lower sign in the interval Δ_{2q}:

$$-q_0+\frac{\gamma_{sp}^2}{\gamma_s} \leqslant q \leqslant q_0-\frac{\gamma_{sp}^2}{\gamma_p}. \quad (15)$$

If Goodwin's condition $\gamma_s\gamma_p = \gamma_{sp}^2$ holds, then $\Delta_{1q} = \Delta_{2q} = 0$. For $\gamma_{sp} = 0$, of course, the upper limit of the lower band coincides with the lower limit of the upper band.

3. GENERAL FORMULATION OF THE PROBLEM OF THE SHOCKLEY SURFACE STATES

Let us consider a finite crystal delimited by a plane perpendicular to the elementary translation $\boldsymbol{\rho}$. The elementary cells lying directly under the surface formed in this way we denote by the index $l = 0$. The perturbation potential V shall have the following properties: $\nu,\nu' = s,p$:

$$\int \bar{w}_{\nu';\eta'\zeta'}(\boldsymbol{r}-l'\boldsymbol{\rho})Vw_{\nu;\eta\zeta}(\boldsymbol{r}-l\boldsymbol{\rho})\,dv = A_{\nu\nu'}\delta_{\eta\eta'}\delta_{\zeta\zeta'}\delta_{l0}\delta_{l'0}, \quad (16)$$

where the one-dimensional Wannier functions are defined as follows:

$$w_{\nu;\eta\zeta}(\boldsymbol{r}-l\boldsymbol{\rho}) = \frac{1}{N}\sum_{mn}\exp[i(\eta m+\zeta n)]a_\nu \times \times(\boldsymbol{r}-l\boldsymbol{\rho}-m\boldsymbol{\sigma}-n\boldsymbol{\tau}). \quad (17)$$

The wave function we assume in the form

$$\psi_{\eta\zeta} = \sum_{l=0}^{P-1}\sum_{\nu=s,p} e_{\nu;\eta\zeta}(l)\,w_{\nu;\eta\zeta}(\boldsymbol{r}-l\boldsymbol{\rho}). \quad (18)$$

P is the number of cells in the direction of the z-axis. In the text we henceforth omit the indices η,ζ.

By specialization of the general equations from Refs. (6) and (7) we obtain the equations for the coefficients appearing in the linear equation (18):

$$e_\nu(l) = \sum_{\nu,\nu'=s,p}[\phi_{\nu\nu'}(l)A_{\nu'\nu''}-\phi_{\nu\nu'}(l+1)\epsilon_{\nu''\nu'}(1)]e_{\nu''}(0). \quad (19)$$

In these equations we have used the following symbols ($N\to\infty$):

$$\phi_{\nu\nu'}(l) = \frac{1}{2\pi}\sum_{x=\mathrm{I,II}}\int_0^{2\pi}\frac{e^{il\xi}\Lambda_{\nu\nu'}^x(\xi)}{W-E^x(\xi)}\,d\xi, \quad (20)$$

$$\epsilon_{\nu\nu'}(l) = \frac{1}{2\pi}\sum_{x=\mathrm{I,II}}\int_0^{2\pi}\Lambda_{\nu\nu'}^x(\xi)E^x(\xi)e^{-il\xi}\,d\xi, \quad (21)$$

$$\Lambda_{ss}^x(\xi) = \frac{1}{1+|\lambda_\xi^x|^2},\ \Lambda_{pp}^x(\xi) = \frac{|\lambda_\xi^x|^2}{1+|\lambda_\xi^x|^2},$$
$$\Lambda_{ps}^x = \overline{\Lambda_{sp}^x} = \frac{\lambda_\xi^x}{1+|\lambda_\xi^x|^2}. \quad (22)$$

In calculating the magnitudes defined by the relations (20) and (21) we utilize the circumstance that we can write the following simple expressions:

$$\Lambda_{ss}^{\mathrm{I}}+\Lambda_{ss}^{\mathrm{II}} = \Lambda_{pp}^{\mathrm{I}}+\Lambda_{pp}^{\mathrm{II}} = 1,\ \Lambda_{sp}^{\mathrm{I}}+\Lambda_{sp}^{\mathrm{II}} = 0,$$
$$\Lambda_{ss}^{\mathrm{II}}-\Lambda_{ss}^{\mathrm{I}} = \Lambda_{pp}^{\mathrm{I}}-\Lambda_{pp}^{\mathrm{II}} = \mathrm{sign}\,(\sin\xi)\frac{B}{\sqrt{(1+B^2)}},$$
$$\Lambda_{sp}^{\mathrm{I}}-\Lambda_{sp}^{\mathrm{II}} = \Lambda_{ps}^{\mathrm{II}}-\Lambda_{ps}^{\mathrm{I}} = -i\,\mathrm{sign}\,(\sin\xi)\frac{1}{\sqrt{(1+B^2}}. \quad (23)$$

Here we have introduced the magnitude

$$B = \frac{q+q_0\cos\xi}{\gamma_{sp}\sin\xi} \quad (24)$$

by means of which we can simply express $\lambda_\xi^{\mathrm{I,II}}$ and $E''(\xi)$:

$$\lambda_\xi^{\mathrm{I,II}} = i\,[B \mp \operatorname{sign}(\sin\xi)\sqrt{(1+B^2)}]$$
$$E''(\xi) = 2|\sin\xi|\gamma_{sp}\sqrt{(1+B^2)}. \qquad (25)$$

Utilizing these relations, we obtain by a simple calculation:

$$\epsilon_{ss}(l) = -2q\delta_{l,0} - \gamma_s(\delta_{l,1}+\delta_{l,-1}),$$
$$\epsilon_{pp}(l) = 2q\delta_{l,0} + \gamma_p(\delta_{l,1}+\delta_{l,-1}),$$
$$\epsilon_{sp}(l) = -\epsilon_{ps}(l) = \gamma_{sp}(\delta_{l,-1}-\delta_{l,1}),$$

$$\phi_{ss}(l) = \frac{1}{2\pi}\int_0^{2\pi} \zeta^{-1}(\xi)[W-2q-2\gamma_p\cos\xi]e^{il\xi}\,d\xi,$$

$$\phi_{pp}(l) = \frac{1}{2\pi}\int_0^{2\pi} \zeta^{-1}(\xi)[W+2q+2\gamma_s\cos\xi]e^{il\xi}\,d\xi,$$

$$\phi_{sp}(l) = -\phi_{ps}(l) = \frac{2i\gamma_{sp}}{2\pi}\int_0^{2\pi} \zeta^{-1}(\xi)\sin\xi e^{il\xi}\,d\xi. \qquad (26)$$

where

$$\zeta(\xi) = W^2 - 4(\gamma_{sp}^2+q^2) + 2(\gamma_{sp}^2-\gamma_s\gamma_p) -$$
$$-2\cos\xi[W(\gamma_p-\gamma_s)+2q(\gamma_s+\gamma_p)] + 2\cos 2\xi \times$$
$$\times(\gamma_{sp}^2-\gamma_s\gamma_p). \qquad (27)$$

Introducing the notation

$$\epsilon' = \gamma_p(W+2q)-\gamma_s(W-2q),\ \epsilon'' = \gamma_s\gamma_p-\gamma_{sp}^2,$$
$$w = W^2-4q^2-2(\gamma_{sp}^2+\gamma_s\gamma_p),$$
$$c = \epsilon'/\epsilon'',\quad x = w/\epsilon''. \qquad (28)$$

we find by comparison with the equations (11) and (3) of Ref. (8) that the integrals at which we arrive in our present paper can be transformed into the integrals solved in publication (8). Therefore holds

$$\phi_{ss}(k) = (W-2q)L_k - \gamma_p(L_{k-1}+L_{k+1}),$$
$$\phi_{pp}(k) = (W+2q)L_k + \gamma_s(L_{k-1}+L_{k+1}),$$
$$\phi_{sp}(k) = -\phi_{ps}(k) = -\phi_{sp}(-k)$$
$$= \gamma_{sp}(L_{k+1}-L_{k-1}), \qquad (29)$$

where

$$L_k = \frac{1}{2\pi}\int_0^{2\pi}\frac{e^{ik\xi}}{\zeta(\xi)}\,d\xi$$
$$= (\epsilon'')^{-1}\frac{1}{(z_4-z_2)(1-z_1z_3)}\left[\frac{z_2^{|k|}}{z_2-z_1} - \frac{z_4^{|k|}}{z_4-z_3}\right] \qquad (30)$$

the magnitudes z satisfy equation (13) of Ref. (8). For this reason, we may in our further calculation also use the relations derived for these magnitudes in the mentioned paper.

The condition for the solubility of the system of equations (19) reads as follows:

$$\begin{vmatrix} \phi_{ss}(0)A_{ss}+\phi_{ss}(1)\gamma_s+\phi_{sp}(1)-1, & \\ \quad \phi_{ss}(0)A_{sp}-\phi_{ss}(1)\gamma_{sp}-\phi_{sp}(1)\gamma_p & \\ \phi_{pp}(0)A_{ps}+\phi_{ps}(1)\gamma_s+\phi_{pp}(1)\gamma_{sp}, & \\ \quad \phi_{pp}(0)A_{pp}-\phi_{ps}(1)\gamma_{sp}-\phi_{pp}(1)\gamma_p-1 & \end{vmatrix} = 0. \qquad (31)$$

4. CASE OF A ZERO SURFACE PERTURBATION POTENTIAL

A characteristic property of the Shockley surface states is the possibility of the existence of these states at a zero perturbing potential. For this reason, we shall consider in this paragraph the case in which we can put

$$A_{ss} = A_{pp} = A_{ps} = 0. \qquad (32)$$

The general relation (31), which defines the energy of the surface states, is then simplified to the form

$$1+\phi_{pp}(1)\gamma_p-\phi_{ss}(1)\gamma_s-2\phi_{sp}(1)\gamma_{sp}$$
$$= +\epsilon''[\phi_{ss}(1)\phi_{pp}(1)-\phi_{sp}^2(1)]. \qquad (33)$$

By substituting the expressions (29) we obtain the following equation:

$$1+\epsilon'L_1+2\epsilon''L_2+2(\epsilon''+2\gamma_{sp}^2)L_0$$
$$= \epsilon''[(W^2-4q^2)L_1^2-\epsilon''(L_0+L_2)^2-4\gamma_{sp}^2L_0L_2-$$
$$-\epsilon'L_1(L_0+L_2)]. \qquad (34)$$

Using the relation (24) from Ref. (8) we can

rearrange equation (34) to the form:

$$R(1-z_1^2z_3^2)\left[2+\frac{4\gamma_{sp}^2}{\epsilon''}-z_2z_4-z_1z_3\right]$$
$$=\epsilon''\left[wL_1^2+2R\epsilon''^{-1}\left(1+\frac{2\gamma_{sp}^2}{\epsilon''}\right)-\right.$$
$$\left.-(L_0+L_2)(\epsilon' L_1+\epsilon'' L_0)\right]. \qquad (35)$$

A short calculation in which we set out from equation (13) of Ref. (8) leads to the final relation, expressed in terms of the magnitudes z:

$$z_1z_3+z_2z_4=\frac{2}{\epsilon''}(1+2\gamma_{sp}^2). \qquad (36)$$

By substitution of the definitions (15) and (16) into equation (36) we obtain

$$6+x+8\,\frac{\gamma_{sp}^2}{\epsilon''}=\text{sign}\,(y_1y_2)[(x-2)^2-4c^2]^{1/2}. \qquad (37)$$

For definition of quantities y_1, y_2 cf. equation (16) of Ref. (8):

We eliminate the root appearing in the preceding formula and after a small manipulation we obtain the expression for the energy of the Shockley surface state at zero perturbation potential (cf. Fig. 1):

$$W=\frac{q(\gamma_s-\gamma_p)}{q_0}. \qquad (38)$$

In order that the signs at the left and right sides of equation (37) be the same, the following inequality must be satisfied:

$$|q|\leqslant q_0 \qquad (39)$$

which represents the condition for the existence of Shockley states at zero perturbation potential.

For our further discussion it is important to determine the wave functions pertaining to the surface states. In the calculation of the coefficients $e_x(l)$ we could set out from equation (19). However, the result is obtained more rapidly, if we substitute the calculated value of the energy of the Shockley surface state into original equations for the coefficients obtained by the usual variation method. We assume

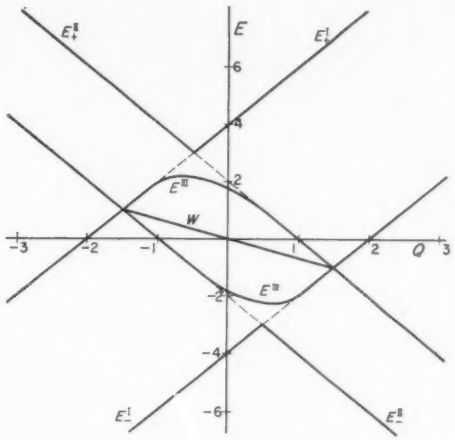


FIG. 1. Dependence of energy bands of volume states and of the energies of the Shockley states W on the quantity q. Cf. equations (12, 13, 38) and definition (8). $\gamma_p=2$, $\gamma_s=\gamma_{sp}=1$.

$$e_s(l)=e_s(0)(\beta_1+\alpha\beta_2),$$
$$e_p(l)=\lambda e_s(l), \qquad (40)$$

where β_1, β_2 are those roots of the equation

$$\frac{\gamma_s-\gamma_p}{2}\left(\frac{q}{q_0}+\beta'\right)=\pm[(q+q_0\beta')^2-\gamma_{sp}^2\beta''^2]^{1/2} \qquad (41)$$

whose absolute values are less than unity.

Equation (41) is an analogy of the relation (7) for the volume states. From equation (41) in which we denote

$$\beta'=\frac{\beta_1+\beta_2}{2},\quad \beta''=\frac{\beta_1-\beta_2}{2}, \qquad (42)$$

we obtain by a simple manipulation

$$\frac{q}{q_0}+\beta'=\pm\left(\frac{\gamma_{sp}}{\gamma_s\gamma_p}\right)^{1/2}\beta''. \qquad (43)$$

Therefore

$$\beta_{1,2}=\frac{1}{1+\sqrt{\left(\frac{\gamma_{sp}^2}{\gamma_s\gamma_p}\right)}}\left[-\frac{q}{q_0}\mp\right.$$
$$\left.\mp\sqrt{\left(\frac{q^2}{q_0^2}+\frac{\gamma_{sp}^2}{\gamma_s\gamma_p}-1\right)}\right]. \quad (44)$$

The equation for the coefficients, following from the variation procedure for the cells within the crystal, is (cf. equation 15):

$$\frac{q}{q_0}=-\beta'+\mu\lambda\beta'', \quad -\frac{q}{q_0}\lambda=-\nu\beta''+\lambda\beta'. \quad (45)$$

Equation (41) is then just the condition for the solubility of the system (45). The conditions for the border cells read:

$$2\frac{q}{q_0}(1+\alpha)=-\beta_1-\alpha\beta_2+\mu\lambda(\beta_1+\alpha\beta_2),$$
$$-2\frac{q}{q_0}\lambda(1+\alpha)=-\nu(\beta_1+\alpha\beta_2)+\lambda(\beta_1+\alpha\beta_2). \quad (46)$$

Hence we simply calculate

$$\lambda=-\sqrt{\frac{\gamma_s}{\gamma_p}}, \quad \alpha=-\frac{\beta_2}{\beta_1} \quad (47)$$

so that we can write the expressions for the coefficients in the form:

$$e_s(l)=e_s(0)\frac{\beta_1^{l+1}-\beta_2^{l+1}}{\beta_1-\beta_2}, \quad \frac{e_p(l)}{e_s(l)}=-\sqrt{\frac{\gamma_s}{\gamma_p}}. \quad (48)$$

finite crystal lie within the band of the permissible energies of a cyclic crystal, to which we approximate the infinite crystal. If the number of cells of the cyclic crystal is very large, the difference between the energy of the electron in the finite crystal and the nearest energies of the electron in the cyclic crystal is very small. The general method of calculating the surface states can also be used for the calculation of the energies of the volume states, however, we must take the integrals, into which the summations pass at a large N, in the sense of an improper integral. The magnitudes ϕ can then be expressed by means of L_k, where (9)

$$L_k=-\frac{1}{2\epsilon'}\frac{\sin|k|\varphi}{\sin\varphi}. \quad (49)$$

The angle φ is defined by the equation

$$2\cos\varphi=\frac{w}{\epsilon'}. \quad (50)$$

The functions L_k for large values of k do not descend to zero, but exhibit a periodic dependence on this parameter. For this reason, it is not possible to consider only one of the borders of the finite crystal, but we must also take into account the other of the limits by which we have singled out the finite crystal from the corresponding infinite crystal. The fundamental equations for the determination of the coefficients read [6–8]:

$$\begin{aligned}
e_s(l)&=[(\phi_{ss}(l+1)\pm\phi_{ss}(P-l))\gamma_s+(\phi_{sp}(l+1)\pm\phi_{sp}(P-l))\gamma_{sp}]l_s(0)-\\
&-[(\phi_{ss}(l+1)\pm\phi_{ss}(P-l))\gamma_{sp}+(\phi_{sp}(l+1)\pm\phi_{sp}(P-l))\gamma_p]l_p(0),\\
e_p(l)&=[(\phi_{ps}(l+1)\mp\phi_{ps}(P-l))\gamma_s+(\phi_{pp}(l+1)\mp\phi_{pp}(P-l))]l_s(0)-\\
&-[(\phi_{ps}(l+1)\mp\phi_{ps}(P-l))\gamma_{sp}+(\phi_{pp}(l+1)\mp\phi_{pp}(P-l))\gamma_p]l_p(0). \quad (51)
\end{aligned}$$

5. VOLUME STATES

For our further considerations it is necessary to show which values of the energies pertaining to the volume states are permitted in the case of a finite crystal consisting of P elementary cells. For reasons of simplicity we restrict ourselves to the case characterized by GOODWIN's[1] condition $\gamma_s\gamma_p=\gamma_{sp}^2$. The energies of the volume states for a

The wave functions of the electron in a finite crystal can be divided into two groups: for symmetric functions

$$e_s(0)=e_s(P-1), \quad e_p(0)=-e_p(P-1) \quad (52a)$$

and for antisymmetric functions

$$e_s(0)=-e_s(P-1), \quad e_p(0)=e_p(P-1). \quad (52b)$$

The condition for the solubility of the system of equations for the coefficients $e_s(0)$ and $e_p(0)$ then are:

$$\epsilon' \sin \phi \pm \bar{\epsilon} \sin P\phi = 0 \qquad (53a)$$

where we have introduced the magnitude

$$\bar{\epsilon} = \gamma_s(W-2q)+\gamma_p(W+2q). \qquad (54)$$

The upper signs apply to symmetric states, the lower signs to antisymmetric states.

For the ratio $e_p(0)/e_s(0)$ it follows that

$$\frac{e_p(0)}{e_s(0)} = \frac{\gamma_p(W+2q)}{\gamma_s(W-2q)}. \qquad (55)$$

The condition for the validity* of

$$e_s(-1) = e_p(-1) = e_s(P) = e_p(P) = 0 \qquad (56)$$

is

$$4\gamma_s\gamma_p+\epsilon' \cos \varphi \pm \bar{\epsilon} \cos(P\psi) = 0 \qquad (53b)$$

The angle φ must simultaneously satisfy the relations (53a) and (33b).

By calculation for an arbitrary special case we find the correctness of SHOCKLEY's statement that the number of states in both the upper and lower bands is decreased by one in the presence of Shockley surface states. We can therefore say that one of the Shockley states originates from the upper and the other one from the lower band of the allowed energies of the infinite crystal.

6. CONCLUSIONS

The condition for the existence of Shockley states, expressed by the inequality (39), is independent of the exchange integral γ_{sp}, which characterizes the interaction of the s and p_z orbitals, and shows that the Shockley states at zero deformation potential are encountered when the perturbation of the levels of the energies pertaining to the s and p_z states in the bands is so considerable that an intersection of the bands has taken place. The energy of the surface state, given by equation (38), is likewise independent of γ_{sp}. It is evident that the lower limit of the upper band and the upper limit of the lower band tend to just this energy value, lying in the gap between the volume state bands, when the parameter γ_{sp} approaches zero. The degree of localization of the wave function of the surface state depends, of course, essentially just on the width of the gap between the bands.

The hybridization in the surface state causes the hybridized atomic orbital for an atom on the surface to have a cloud of high electron density projecting out of the crystal. On the other hand, the hybridized orbitals of the surface atoms ($l = 0$) of the lower occupied band have a decreased electron density at the crystal surface and an increased density in places where an overlap can occur with the electron cloud of an atom in the second layer ($l = 1$) from the surface.

The wave functions of the surface states of the lower band, in terms of the band theory, characterize the states of the electrons produced by a bond between the atoms, while the surface states of the Shockley type in our model characterize the unsaturated valences of the surface atoms projecting out of the crystal. It seems that in general the Shockley surface states are associated with the special bond conditions of the crystal surface and equivalent to the chemical concept of a free valence of the surface.

* Cf. equation (15b) of Ref. 7.

REFERENCES

1. GOODWIN E. T., *Proc. Camb. Phil. Soc.* **35,** 232 (1939).
2. ARTMANN K., *Z. Phys.* **131,** 244 (1952).
3. TAMM Ig., *Phys. Z. Sowjet.* **1,** 733 (1932); *Z. Phys.* **76,** 849 (1932).
4. SHOCKLEY W., *Phys. Rev.* **56,** 317 (1939).
5. LIPMANN B. A., *Ann. Phys., N.Y.* **2,** 16 (1957).
6. KOUTECKÝ J., *Phys. Rev.* **108,** 13 (1957).
7. KOUTECKÝ J. and FINGERLAND A., *Dokl. Akad. Nauk SSSR* **125,** 841 (1959).
8. TOMÁŠEK M. and KOUTECKÝ J., *Czech. J. Phys.*, in press.
9. KOUTECKY J. and FINGERLAND A., *Coll. Czech. Chem. Commun.* **25,** 1 (1960).

VOL. 14 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 248–267. Printed in Great Britain.

INFLUENCE OF TRAPPING, DIFFUSION AND RECOMBINATION ON CARRIER CONCENTRATION FLUCTUATIONS

M. LAX and P. MENGERT*

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

Abstract—Carrier concentration fluctuations are found to be relatively large in regions of low concentration. Diffusion is shown to add a $(1/\omega)^{3/2}$ boundary term (important at high frequencies) to the usual volume recombination spectrum. This result is found to be valid independent of the geometry or the number of dimensions. But for the case of surface recombination it is multiplied by a frequency dependent probability that an electron having left the region will not return. Traps of large cross section are characterized by a fast trapping rate and an exceedingly slow effective recombination rate caused by multiple trapping. The fast rate is unaffected by diffusion. The slow rate is shown by a method of low mode isolation to be approximately describable in terms of an effective diffusion constant reduced in about the same ratio as the recombination. The resulting $(1/\omega)^{3/2}$ spectrum is thus shifted to a low frequency region and increased enormously in strength.

1. INTRODUCTION AND SUMMARY

ALTHOUGH this paper has been prepared for presentation at the Second Conference on Semiconductor Surfaces, it may enjoy the distinction of being the only paper that is not, strictly speaking, about surface properties *per se*. It is being presented here because surface physics people are interested in noise, which is known to originate largely near the surface, and because the general methods developed here can be applied to surface problems, as well as to the bulk.

The most challenging problem is, of course, the $1/f$ noise problem.[1] Experimental evidence exists which favors a surface origin for this noise.[2] There also is more or less general agreement that $1/f$ noise is associated with carrier concentration fluctuations.[3] Usually concentration fluctuations will be most significant if the conductivity is controlled by the number of carriers in some small region, since then, the percentage fluctuations will be greatest. Space-charge regions near the surface may provide such small regions, particularly if an inversion layer is present, since the required intrinsic region will contain a small number of carriers.[4]

Among the factors which may contribute to $1/f$ noise are trapping, drift, diffusion, recombination and fluctuations in space charge, surface potential, barrier heights, and tunneling. One of the difficulties of explaining $1/f$ noise over a wide frequency range is that it requires a broad distribution of relaxation times, extending down to very low frequencies[5] (as low as 10^{-4} c/s). Models have in fact been proposed[6] which yield $1/f$ noise by introducing in a more or less *ad hoc* way, the appropriate distribution of relaxation times. A satisfactory explanation requires, however, that one show why such a distribution of relaxation times is a natural consequence of more elementary properties of the system.

Although they do not in themselves lead to $1/f$ noise, this paper will consider in detail voltage fluctuations induced by carrier concentration fluctuations in a homogeneous bulk semiconductor as modified by *diffusion* and *trapping*—specifically multiple trapping of the type observed by HORNBECK and HAYNES[7] the former because it automatically converts individual relaxation times into a distribution of relaxation times; the latter because it can lead to extremely long relaxation times, e.g. 10^6 sec if the concentration of traps is $10^{16}/\text{cm}^3$ (which could easily occur in a surface region).

* Now at Physics Department, Brandeis University, Waltham, Mass.

VOL. 14 1960

Specifically we propose to show that:

(1) An intrinsic region can lead to large fluctuations in carrier concentration.

(2) Diffusion will modify the volume recombination noise by adding a boundary term with $(1/\omega)^{3/2}$ spectrum at high frequencies regardless of geometry or number of dimensions.

(3) Deep traps with a trapping rate ν, large compared to the recombination rate r, and a long "sitting time" $1/g'$ lead to multiple trapping (ν/r times) and an effective lifetime $(\nu/r)(1/g')$ which in the HORNBECK–HAYNES experiments[7] was several hundred seconds long. See Table 1 for numerical values.

(4) The total noise associated with traps and transport (diffusion and drift) can be written as a volume term associated with fluctuations in the total number of carriers plus a term associated with transport of carriers across the boundary.

(5) The fast rate $\lambda_F \approx \nu$ at which electrons disappear from the conduction band is essentially unmodified by transport. Hence the total noise is the usual volume term plus the transport modification of the slow modes (which at infinite wavelength have the slow decay rate $\lambda_S \approx rg'/\nu$ associated with multiple trapping).

(6) When the density of free electrons n is small compared to the capture concentration[8], $\hat{n}' = \hat{n}[1-\hat{n}/n_t]$ = number of trapped electrons × fraction of empty traps, the noise comes predominantly from the slow modes and these yield conductivity fluctuations associated primarily with fluctuations in the trapped carrier density.

(7) Although a precise method for dealing with the coupled transport of free and bound carriers is developed (Appendix E) one can with phenomenal accuracy use a method of low mode isolation: the free carriers because of their fast decay rate are assumed to adjust to the instantaneous concentration of trapped carriers, i.e. the trapped carrier density alone is to a good approximation a Markoffian variable.

Volume effects

Because of charge neutrality in the form $\Delta P = \Delta N + \Delta\hat{N}$ concentration fluctuations lead to conductivity fluctuations, and hence to voltage fluctuations of the form[9] (see equation (2.9)):

$$\Delta V = [V/(P+Nb)][\Delta\hat{N}+(1+b)\Delta N] \quad (1.1)$$

where $N, \hat{N}$ and P are the total number of free electrons, bound electrons and holes, respectively, and b is the electron/hole mobility ratio. The integrated noise is then (see equation (2.20)) given by

$$\int_0^\infty G(V,f)\,df = \langle(\Delta V)^2\rangle$$

$$= [V^2/(P+Nb)^2][\hat{N}'+(1+b)^2N] \quad (1.2)$$

where $f = \omega/2\pi$ and

$$\hat{N}' = \hat{N}[1-\hat{N}/N_t] \quad (1.3)$$

is the total number of bound electrons times the fraction of empty traps. This result, equation (1.2), is clearly largest in magnitude in the intrinsic region $N \sim P$. For $N \ll \hat{N}'$ the noise comes primarily from the trapped carriers, see above remark (6).

The volume noise associated with ΔN and $\Delta\hat{N}$ can be computed exactly (equation (2.22)) and decomposed into a slow and fast contribution with the approximate result (compare with equation (2.29)):

$$G(V,f) \approx \frac{4V^2}{(P+Nb)^2}\left[\frac{\hat{N}'\lambda_S}{\omega^2+\lambda_S^2}+(b^2+b)N\frac{\lambda_F}{\omega^2+\lambda_F^2}\right] \quad (1.4)$$

where the fast and slow eigenvalues associated with trapping and release of carriers are given, more accurately than before as

$$\lambda_F \approx \nu+r+g'; \quad \lambda_S = (rg')/\lambda_F \quad (1.5)$$

Diffusion plus recombination, no traps

Our results on diffusion and recombination *without trapping* can be written, using equation (1.1) with $\Delta\hat{N} = 0$ in the form:

$$G(V,f) = 4[V^2(1+b)^2/(P+Nb)^2]\langle(\Delta N)^2\rangle g(\omega) \quad (1.6)$$

where $\langle(\Delta N)^2\rangle = NP/(N+P)$ according to equation (B12) and

$$g(\omega) = Re\int_0^\infty \exp(-i\omega t)\Phi(t)\,dt \tag{1.7}$$

and

$$\Phi(t) \equiv \langle\Delta N(t)\Delta N(0)\rangle/\langle(\Delta N)^2\rangle \tag{1.8}$$

$$= \frac{1}{\Omega}\int_\Omega d\mathbf{r}\int_\Omega d\mathbf{r}' K(\mathbf{r},\mathbf{r}',t) \tag{1.9}$$

with Ω the one-, two- or three-dimensional domain containing the carriers and $K(\mathbf{r},\mathbf{r}',t)$ represents the Green's function of the transport equation, i.e. the probability that an electron will be found at $\mathbf{r}$ at time t if it was known to be at $\mathbf{r}'$ at time 0. Thus $\Phi(t)$ represents the probability that an electron initially in Ω with uniform probability $d\mathbf{r}'/\Omega$ will still be in Ω at time t, i.e. $\Phi(t)$ is the *photoconductive decay* function after an initial uniform excitation. It is also referred to as the Smoluchowski after-effect function.[10]

Boundary effects

The effect of the recombination rate r on $\Phi(t) \equiv \Phi(r,t)$ is simply

$$\Phi(r,t) = \exp(-rt)\Phi(0,t) \tag{1.10}$$

so that $g(\omega)$ is a function of $(r+i\omega)$ whose properties depend on diffusion (and drift) across the boundary. The asymptotic properties of $g(\omega)$ for large ω depend[11] on $\Phi(t)$ for small t. Since the diffusive motion for small t is of order $(Dt)^{1/2}$ where D is the diffusion constant, we conjecture that the particles within a distance $(Dt)^{1/2}$ of the surface escape. (See Fig. 1.) The fraction escaping

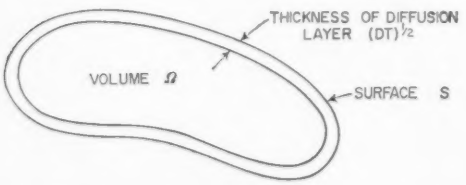


FIG. 1. Diffusion at the boundary. In time t particles in the diffusion region $(Dt)^{1/2}\times$(surface area S) strike the surface. The fraction of the total number of particles in volume Ω which escape is then $(Dt)^{1/2}(S/\Omega)Q$ where Q is the probability that a particle which strikes the surface does not return.

in small time t is then proportional to $(Dt)^{1/2}S/\Omega$ and we can write

$$\Phi(0,t) \approx 1-Q(Dt)^{1/2}S/\Omega \tag{1.11}$$

where Q is an as yet undetermined parameter of order of magnitude unity, S is the generalized surface and Ω the generalized volume. For example for a sphere

$$S/\Omega = 4\pi R^2/(4\pi R^3/3) = 3/R \tag{1.12}$$

For a rectangle

$$S/\Omega = 2(L_1+L_2)/L_1L_2 \tag{1.13}$$

for a circle

$$S/\Omega = 2\pi R/\pi R^2 = 2/R \tag{1.14}$$

for a line of length L in one dimension

$$S/\Omega = 2/L \tag{1.15}$$

because there are two boundary points.

Universal inverse three-halves power law

In summary we expect that at all frequencies

$$g(\omega) = Re(r+i\omega)^{-1}+\Delta g(\omega) \tag{1.16}$$

when the first term is the volume term and the second term is the boundary term. At high frequencies (1.11) implies that the boundary term is:

$$\Delta g(\omega) \approx -(S/\Omega)Re\{[D/(r+i\omega)^3]^{1/2}Q\} \tag{1.17}$$

or, at high frequencies, when Q is real we expect an inverse three-halves power law:

$$\Delta g(\omega) \approx (S/\Omega)Q(D/2\omega^3)^{1/2} \tag{1.18}$$

Our more detailed investigation of the above mentioned one-, two- and three-dimensional examples suggests that the as yet undetermined parameter Q can be interpreted as the probability that an electron having reached the boundary will not return. For a "fictitious" surface drawn in an otherwise homogeneous medium, we find, regardless of the number of dimensions or shape

$$Q = \tfrac{1}{2} \tag{1.19}$$

If the outside region has recombination constant r', and diffusion constant D', then we find by using a one-dimensional example that

$$Q = \gamma'/(\gamma'+\gamma) \xrightarrow[\omega\to\infty]{} \sqrt{D}/(\sqrt{D}+\sqrt{D'}) \tag{1.20}$$

where

$$\gamma = [(r+i\omega)/D]^{1/2}; \quad \gamma' = [(r'+i\omega)/D']^{1/2} \quad (1.21)$$

In the same way we find that for surface recombination with velocity s, the boundary condition at $x = L$

$$-D \text{ grad } n = sn \quad (1.22)$$

leads to

$$Q = (s/\gamma D)/[1+(s/\gamma D)] \quad (1.23)$$

which is close to unity for large surface recombination, except as $\omega \to \infty$ where Q varies as $\omega^{-1/2}$. Also, if we terminate with the "characteristic impedance" $s = \gamma D$, Q reduces to its value, $\frac{1}{2}$, for the infinite case.

VOL. 14 1960

Boundary interference effects

Equation (1.17) turns out to be *exactly correct* in one-dimension if we include in Q interference effects between the two boundaries. For a line of length $L = 2B$, with drift neglected, we find in Sections 3 and 4 that "fictitious" surfaces lead to

$$Q = \tfrac{1}{2}[1 - \exp(-2\gamma B)] \quad (1.24)$$

whereas the correct result with surface recombination at $x = \pm B$ is shown in Section 4 to be

$$Q = (s/\gamma D)/[1+(s/\gamma D)\coth(\gamma B)] \quad (1.25)$$

in agreement with results obtained by CHAMPLIN[12] using a Langevin approach involving some *ad hoc* but evidently sound assumptions* about the noise sources based on a transmission line analogy. Boundary interference effects can be made to disappear by letting $B \to \infty$, in fact, equations (1.19) and (1.23) were deduced in this way from the computed results, equations (1.24) and (1.25). Our interpretation was also checked by solving exactly an example in which the surface recombination velocity had different values at the two ends of the specimen, see equation (4.20), and by considering a semi-infinite example.

Trapping, recombination and diffusion

For "fictitious" surfaces, the above results were computed making use of the fact that a plane wave $\exp(ikx)$ decays as $\exp[-\lambda(k)t]$ with

$$\lambda(k) = r + Dk^2. \quad (1.26)$$

* CHAMPLIN, however, states emphatically that correlations in density fluctuations at one time must occur over distances of the order of a diffusion length in contradiction to our equations (3.5) and (C1).

When traps are present, our technique of low mode isolation leads to the slow decay of a plane wave of trapped carriers with

$$\lambda_S(k) \approx (g'r + Dv_A k^2)/(\nu + r + Dk^2) \quad (1.27)$$

where

$$v_A = g' + \nu n/(n+p)$$

(drift is omitted here, but discussed in Section 5), so that $\lambda_S(0) = \lambda_S$ is the effective volume decay rate, in agreement with equation (1.5). For $\omega < \nu_A$, we can set $k = 0$ in the denominator, obtaining effective recombination and diffusion constants:

$$r_{\text{eff}} = \lambda_S \approx r[g'/(\nu+r)] \quad (1.28)$$

$$D_{\text{eff}} = D[\nu_A/(\nu+r)] \quad (1.29)$$

and the results of our previous discussion of recombination plus diffusion (without trapping) apply with these effective constants, which are about 10^{-7} times the usual constants. The $(1/\omega)^{3/2}$ behavior is indeed much more readily observable with trapping than without. Using equation (1.18) with $SQ/\Omega = 1/L$, and r/ω^2 for the volume noise, we find that the diffusion contribution exceeds the volume contribution at frequencies greater than

$$\omega = 2r^2L^2/D = 2r(L/\mathscr{L}_d)^2 \quad (1.30)$$

where $\mathscr{L}_d = (D/r)^{1/2}$, the diffusion length, is usually small compared to L. At this high frequency, the noise is only

$$\Delta g = D^2/(4r^3L^4) \quad (1.31)$$

and may be difficult to observe in the presence of thermal or other sources of noise. If, however, because of trapping, one inserts effective diffusion and recombination constants 10^7 times smaller, this noise is increased by a factor 10^7 and can probably be observed.

A more careful analysis of equation (1.27), without using the approximation $k = 0$ in the denominator yields the correct diffusion modification of the slow modes. The total volume noise is given by equation (1.4) with $\lambda_S/(\omega^2+\lambda_S^2)$ replaced by

$$g(\omega) = \frac{\lambda_S}{\lambda_S^2+\omega^2} - \frac{1}{L}\left(\frac{D}{\nu+r}\right)^{1/2} \times$$

$$\times Re\frac{\nu_A-\lambda_S}{[(\nu_A+i\omega)(\lambda_S+i\omega)^3]^{1/2}} \quad (1.32)$$

which reduces to our previous result (1.17) with

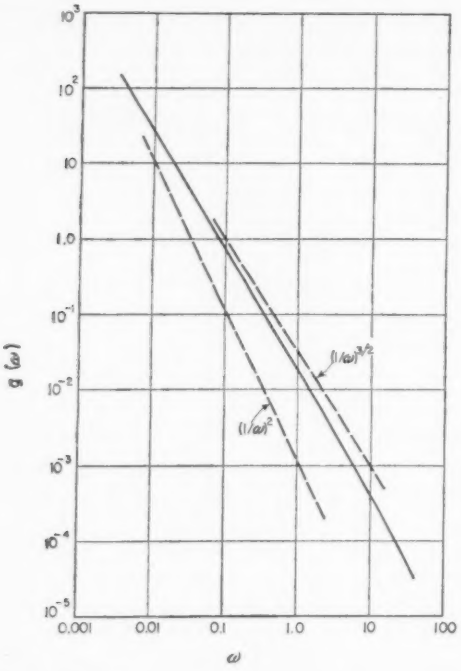


FIG. 2. The voltage noise spectrum at frequency f in the presence of diffusion and trapping is written in the form

$$G(V,f) = 4[V^2\hat{N}'/(P+Nb)^2]g(\omega)$$

and $g(\omega)$ is plotted against $\omega = 2\pi f$. Here V is the applied voltage, P the number of holes, N the number of electrons, b the electron–hole mobility ratio and $\hat{N}'$ is the number of trapped electrons times the fraction of empty traps.

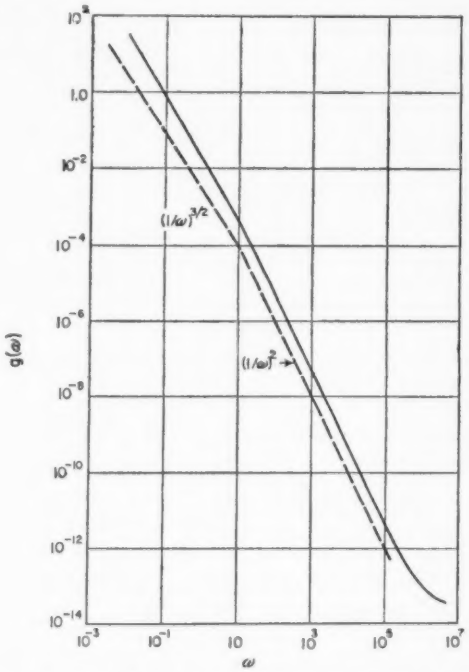


FIG. 3. The voltage noise spectrum as in Fig. 2 plotted over a wider frequency range. The contribution of the fast trapping rate begins to be significant near $\omega = 10^6$ where a level off starts.

effective constants when $\omega \ll \nu_A$ (and $\lambda_S \ll \nu_A$) but reduces to an inverse square law

$$\Delta g \approx D^{1/2}\nu_A/[(\nu+r)^{1/2}L\omega^2] \qquad (1.33)$$

for $\omega \gg \nu_A$. (See Fig. 2.) This diffusion modification of the slow modes will dominate the fast mode contribution (second term of equation (1.4)) up to

$$\omega = [(D\nu)^{1/2}\nu_A\hat{N}'/(NL)]^{1/2} \approx 10^6/\text{sec} \quad (1.34)$$

using the values in Table 1. This behavior is shown in Fig. 3, where one can also see the beginning of the level off associated with the contribution of the fast modes for $\omega \ll \lambda_F \approx \nu$.

We did not expect that our treatment of diffusion, drift and multiple trapping would explain $1/f$ noise. Additional mechanisms that should be considered include:

(1) The fact that a surface region is not homogeneous and possesses space charge which fluctuates.

(2) Traps, recombination centers, donors and acceptors are not uniformly distributed except on the average, and fluctuations can be introduced in the time of travel between centers.

(3) Tunneling between localized centers will fluctuate widely with the separation between nearest neighbors.

(4) Fluctuations may occur in tunneling through barriers that might exist between the surface region and the bulk.[6]

We should emphasize, however, that if the concentration of traps in the surface region is much larger than in the bulk, as is reasonable, the

Table 1. *Typical numerical values of parameters in* HORNBECK–HAYNES[7] *multiple trapping experiments*

Experimental Constants	
Trapping rate	$\nu \sim 3 \times 10^7$/sec.
Recombination rate	$r \sim 3 \times 10^4$/sec.
Effective release rate	$g' \sim 3$/sec.
Majority carrier density	$p \sim 10^{14}$ holes/cm^3
Minority carrier density*	$n \sim 10^7$ electrons/cm^3
Capture concentration†	$\hat{n}' \sim 10^{13}$ electrons/cm^3
Electron diffusivity	$D_n \sim 30$ cm^2/V sec.
Mobility ratio	$b = \mu_n/\mu_p \sim 1200/500 = 2{\cdot}4$
Distance between electrodes	$L = 0{\cdot}1$ cm (our choice)
Derived Constants	
Slow decay rate	$\lambda_S \approx rg'/\nu \sim 3 \times 10^{-3}$/sec.
Effective trapped carrier diffusivity	$D_{\text{eff}} = D\nu_A/(\nu+r) \approx 6 \times 10^{-6}$ cm^2/sec.
Lower limit of 3/2 power law	$2\lambda_S^2 L^2/D_{\text{eff}} \approx 3 \times 10^{-2}$/sec.
Upper limit, 3/2 power law	$\nu_A = g' + \nu n/(n+p) \approx 6$/sec.
Frequency at which fast modes start to contribute	$[(D\nu)^{1/2}\nu_A n'/nL]^{1/2} \approx 10^6$/sec.

* This value, and all values of minority carrier density stated by HORNBECK and HAYNES[7] should be reduced by more than a factor 10 since the *np* product for silicon is now known to be about $0{\cdot}6 \times 10^{20}$ at room temperature based on improved knowledge of the density of states[13] and on the more accurate value 1·117 eV of the energy gap[14] which is 0·077 eV larger than the value used by HORNBECK and HAYNES.

† Capture concentration = (trap concentration × fraction empty) × fraction full. Only the number of empty traps was measured but we assume here that the capture concentration did not differ significantly from this. See however footnote § on p. 577 of Ref. 8.

multiple trapping ratio ν/r could be greatly increased over the value used in this paper and it would be easy to get noise which increases down to 10^{-6} or 10^{-7} c/s.

2. VOLUME EFFECTS: FLUCTUATIONS IN TOTAL NUMBERS OF FREE AND BOUND CARRIERS

Assuming the voltage fluctuation $\Delta V(t)$ is a stationary random variable, the (squared) voltage noise spectrum can, by the Wiener–Khintchin theorem be written in the form[9]

$$G(V,f) = 4Re \int_0^\infty \exp(-i\omega t)\langle \Delta V(t)\Delta V(0)\rangle \, dt \tag{2.1}$$

(cf. Ref. 1, equation 2.6 ff.). If ΔV can be expressed

$$\Delta V = \sum C_r \alpha_r \tag{2.2}$$

as a linear combination of variables α_r, which constitute a Markoffian set $\boldsymbol{\alpha}$ obeying the phenomenological equation

$$d\langle \boldsymbol{\alpha} \rangle / dt + \boldsymbol{\Lambda} \langle \boldsymbol{\alpha} \rangle = 0 \tag{2.3}$$

with the decay matrix $\boldsymbol{\Lambda}$, then

$$G(V,f) = 2Re \sum C_r C_s G^+_{rs} \tag{2.4}$$

where

$$G^+_{rs} = 2\int_0^\infty \exp(-i\omega t)\langle \alpha_r(t)\alpha_s(0)\rangle \, dt$$

$$= 2\sum (i\omega + \boldsymbol{\Lambda})^{-1}_{rj} \langle \alpha_j \alpha_s \rangle \tag{2.5}$$

or

$$G(V,f) = 4Re \sum C_r (i\omega + \boldsymbol{\Lambda})^{-1}_{rj} \langle \alpha_j \alpha_s \rangle C_s \tag{2.6}$$

where summation over repeated indices is understood and ω is written briefly for ω times the unit matrix (cf. Ref. 9, equation (3.13)).

In the *absence* of drift, diffusion or surface recombination, the total number of free and bound carriers predict their own future, and we can choose as our Markoffian set

$$\alpha_1 = \Delta N \qquad \alpha_2 = \Delta \hat{N} \tag{2.7}$$

(whereas if any of these effects are *present*, one must

use the doubly infinite set of variables $\Delta n(x)$ and $\Delta n(x)$ since the future even of the total number of particles then depends on their present spatial distribution). The hole density does not fluctuate independently because of the space-charge neutrality condition

$$\Delta P = \Delta N + \Delta \hat{N} \tag{2.8}$$

which can be expected to be satisfied at frequencies small compared to the dielectric relaxation time. See, however, Appendix C.

For geometries such that one can set up a uniform d.c. current I, the voltage fluctuation ΔV induced by a resistance fluctuation ΔR at constant I obeys

$$\frac{\Delta V}{V} = \frac{\Delta R}{R} = -\frac{\Delta P \mu_p + \Delta N \mu_n}{P\mu_p + N\mu_n} \tag{2.9}$$

Inserting charge neutrality, equation (2.8), we retrieve our starting expression (1.1) for ΔV in terms of ΔN and $\Delta \hat{N}$ with the respective coefficients

$$C_1 = (1+b)C_2; \qquad C_2 = V/(P+Nb) \tag{2.10}$$

where P and N are the steady-state number of holes and electrons and $b = \mu_n/\mu_p$ is the mobility ratio.

The decay matrix $\mathbf{\Lambda}$ must be found by writing down the phenomenological equations (cf. Appendix A) and linearizing them about the steady-state:

$$\begin{aligned} d\langle N\rangle/dt &= \langle G - R + g\hat{N} - (N_t - \hat{N})NC/\Omega\rangle \\ d\langle \hat{N}\rangle/dt &= \langle (N_t - \hat{N})NC/\Omega - g\hat{N}\rangle \end{aligned} \tag{2.11}$$

where R is the rate of recombination of electrons with holes, G is the thermal (and possibly also optical) rate of generation of electron–hole pairs, g is the rate of generation out of the traps, C/Ω is the rate of capture per empty trap, for one electron in the volume Ω and N is the total number of traps. In equation (2.11) alone, N and $\hat{N}$ represent the instantaneous number of free and bound electrons respectively. But we shall immediately linearize $N \to N + \Delta N, \hat{N} \to \hat{N} + \Delta \hat{N}$ and keep $N, \hat{N}$ for the steady-state values. The linearized equations take the form:

$$\frac{d}{dt}\begin{pmatrix}\langle \Delta N\rangle \\ \langle \Delta \hat{N}\rangle\end{pmatrix} = -\mathbf{\Lambda}\begin{pmatrix}\langle \Delta N\rangle \\ \langle \Delta \hat{N}\rangle\end{pmatrix} \tag{2.12}$$

where

$$\mathbf{\Lambda} = \begin{pmatrix} \nu + r, & -g' \\ -\nu, & g' \end{pmatrix} \tag{2.13}$$

and the steady-state values chosen to make the right-hand side of equation (2.11) vanish obey:

$$\begin{aligned} g\hat{N} &= C(N_t - \hat{N})N/\Omega \\ R &= G \end{aligned} \tag{2.14}$$

In extrinsic p-type material, the effect of Shockley–Read–Hall recombination centers can be represented away from the steady-state as

$$R - G \approx r\Delta N \tag{2.15}$$

where r is the recombination rate. (In equation (2.15) a small term proportional to ΔP has been neglected, but could easily have been retained.) The capture rate ν is defined by

$$\nu = C(N_t - \hat{N})/\Omega = C(n_t - \hat{n}) \tag{2.16}$$

where n_t and $\hat{n}$ are the trap density and steady-state density of trapped electrons. (C represents a capture cross section times an electron velocity.) The effective release rate g' is defined by

$$g' = g + CN/\Omega = g + Cn. \tag{2.17}$$

(HORNBECK and HAYNES[7] arrived at the linearized equations (2.13) without obtaining the replacement of g by g'. Their quoted measured values of g are in reality measurements of g', but the correction is usually negligible, see Table 1.)

To evaluate the noise via equation (2.6), we need the reciprocal matrix element

$$(i\omega + \mathbf{\Lambda})^{-1} = \frac{1}{M}\begin{pmatrix} g' + i\omega, & g' \\ \nu, & \nu + r + i\omega \end{pmatrix} \tag{2.18}$$

where

$$M = \det(i\omega + \mathbf{\Lambda}) = rg' + i\omega(\nu + r + g') - \omega^2 \tag{2.19}$$

and we need the fluctuation moments at a given time. *If* we could assume that the free and bound carriers are separately in thermal equilibrium with reservoirs (possibly at different chemical potentials) then we could write

$$\begin{aligned} \langle (\Delta N)^2\rangle &= N \\ \langle (\Delta \hat{N})^2\rangle &= \hat{N}[1 - (\hat{N}/N_t)] \equiv \hat{N}' \\ \langle \Delta N \Delta \hat{N}\rangle &= 0 \end{aligned} \tag{2.20}$$

VOL. 14 1960

in view of Ref. (9), equations (4.25), (4.27) or—for the nonequilibrium steady-state—(12.28). Since trapped electrons interact only with the free electrons and not with a separate reservoir, equations (2.20) are not obviously valid. We justify them by a kinetic derivation in Appendix A, whereas in Appendix B, a more "thermodynamic" argument with explicit consideration of the holes indicates that equation (2.20) is valid when $P \gg N$, $P \gg \hat{N}'$, i.e. when enough holes are present to act as a suitable reservoir.

By inserting equations (2.10), (2.18) and (2.20) into equation (2.6) the noise is given by:

$$\begin{aligned} G(V,f) &= 4C_2^2 Re\{(i\omega+\mathbf{\Lambda})^{-1}_{22}\hat{N}' + \\ &+(1+b)[(i\omega+\mathbf{\Lambda})^{-1}_{12}\hat{N}'+(i\omega+\mathbf{\Lambda})^{-1}_{21}N]+ \\ &+(1+b)^2(i\omega+\mathbf{\Lambda})^{-1}_{11}N\} \\ &= 4C_2^2 Re\{[(\nu+r+i\omega)\hat{N}'+(1+b)(\hat{N}'g'+N\nu)+ \\ &+(1+b)^2N(g'+i\omega)]/M\} \end{aligned} \tag{2.21}$$

or

$$G(V,f) = \frac{4V^2}{(P+Nb)^2}\,\frac{f_0+f_2\omega^2}{(rg'-\omega^2)^2+\omega^2(\nu+r+g')^2} \tag{2.22}$$

where

$$\begin{aligned} f_0 &= rg'(\nu+r)\hat{N}'+(1+b)rg'(g'\hat{N}'+\nu N)+ \\ &+(1+b)^2r(g')^2N \end{aligned} \tag{2.23}$$

$$f_2 = (b^2+b)\nu N - bg'N' + (1+b)^2 rN \tag{22.4}$$

The accuracy of our work may be verified by comparing the integrated noise with equation (1.2). This integration, and an understanding of the nature of the system can best be obtained by making a partial fraction expansion of equation (2.22) in the form:

$$G(V,f) = \frac{4V^2}{(P+Nb)^2}\left[\frac{A_S\lambda_S}{\omega^2+\lambda_S^2}+\frac{A_F\lambda_F}{\omega^2+\lambda_F^2}\right] \tag{2.25}$$

$$\begin{aligned} A_S &= [(f_0/\lambda_S)-f_2\lambda_S]/(\lambda_F^2-\lambda_S^2) \\ A_F &= [f_2\lambda_F-(f_0/\lambda_F)]/(\lambda_F^2-\lambda_S^2) \end{aligned} \tag{2.26}$$

where λ_S and λ_F are the roots of the denominator M corresponding to the slow and fast eigenvalues of $\mathbf{\Lambda}$. The trace and determinant of $\mathbf{\Lambda}$ yield the relations

$$\begin{aligned} \lambda_F &= \nu+r+g'-\lambda_S \\ \lambda_S &= rg'/\lambda_F \end{aligned} \tag{2.27}$$

Since we assume $\lambda_S \ll \lambda_F$, let us iterate, starting with $\lambda_S = 0$. The first approximation

$$\begin{aligned} \lambda_F &\approx \nu+r+g' \\ \lambda_S &\approx rg'/(\nu+r+g') \end{aligned} \tag{2.28}$$

for values appropriate to a typical Hornbeck–Haynes sample (see Table 1) yields $\lambda_F \approx \nu \approx 3\times10^7$, $\lambda_S \approx rg'/\nu \approx 3\times10^{-3}$ so that our relative error is $\lambda_S/\lambda_F \approx 10^{-10}$! In view of the enormous disparity of frequencies equation (2.26) can be adequately approximated by

$$\begin{aligned} A_S &\approx \hat{N}'+(1+b)N \\ A_F &\approx (b^2+b)N \end{aligned} \tag{2.29}$$

without disturbing the integrated noise. The part of the noise associated with fluctuations in hole conductivity alone, can be obtained by setting $b = 0$. However, hole fluctuations are generated by fluctuations in the number of free as well as bound electrons. The influence of the latter alone can be obtained by setting $b = -1$ leaving $A_S \approx \hat{N}'$, $A_F \approx 0$. Since in the Hornbeck–Haynes case $N \ll \hat{N}'$, we see that the predominant contribution to the noise comes from hole fluctuations required to neutralize the (slow) fluctuations of trapped electrons. The slow mode of decay is, to excellent approximation (error g'/ν) just the decay of trapped electrons, whereas the fast mode with the same accuracy corresponds to the decay of free electrons and produces noise proportional to the free electron fluctuations.

The low frequency contribution to equation (2.25) or (1.4) can now be seen to dominate the high frequency contribution up to the fairly high dividing frequency

$$\omega_d \approx (rg'\hat{N}'/N)^{1/2} \approx 3\times10^5/\text{sec} \tag{2.30}$$

(In the presence of diffusion, the dividing frequency is even higher, see equation (5.35) and Fig. 3.) At this high frequency, the noise is enormously reduced and probably unobservable because of other sources of noise. It would be convenient, especially in more complicated problems (e.g. involving diffusion), to be able to calculate the slow mode

contribution directly. This elimination of the fast modes can be accomplished by an adiabatic approximation which states that the fast variable(s) (the free electron concentration) adjusts to the instantaneous value of the slow variable (the trapped electron concentration) as if the latter were fixed. Thus we set $d\langle \Delta N\rangle/dt = 0$ in equation (2.12) and obtain

$$\langle \Delta N\rangle = [g'/(\nu+r)]\langle \Delta \hat{N}\rangle \tag{2.31}$$

The second equation of (2.12) then yields

$$d\langle \Delta \hat{N}\rangle/dt \approx -\lambda_S \langle \Delta \hat{N}\rangle \tag{2.32}$$

with

$$\lambda_S \approx g'r/(\nu+r) \tag{2.33}$$

which agrees, within one part in 10^7 with the more precise answer, equation (2.28). The resulting noise

$$G(V,f) \approx \frac{4V^2}{(P+Nb)^2}\frac{\hat{N}'\lambda_S}{\omega^2+\lambda_S^2} \tag{2.34}$$

is identical to the first term of equation (1.4) and in good agreement with equation (2.29) when $N \ll \hat{N}'$. Since the NP product is fixed, this noise can be maximized if $P = Nb$, i.e. in the intrinsic region. This conclusion agrees with MacRae's[4] experimental finding: $1/f$ noise is strong when an inversion layer (and hence an intrinsic region) is present near the surface.

3. DIFFUSION AND RECOMBINATION WITHOUT TRAPPING: A UNIVERSAL INVERSE THREE-HALVES POWER LAW AT HIGH FREQUENCIES

In the absence of trapping $\Delta\hat{N} = 0$ and the noise is described by equations (1.6)–(1.8). Since $\Phi(0) = 1$, the spectral distribution $g(\omega)$ is normalized to

$$\int_0^\infty g(\omega)\,df = \tfrac{1}{4} \tag{3.1}$$

For volume recombination the after-effect function $\Phi(t)$ is a simple exponential $\exp(-rt)$. In the presence of drift and/or diffusion ΔN is no longer a Markoffian variable but can be expressed in terms of the density which is:

$$\Delta N(t) = \int_\Omega \Delta n(\mathbf{r},t)\,d\mathbf{r} \tag{3.2}$$

$$\langle \Delta N(t)\Delta N(0)\rangle = \int_\Omega\int_\Omega d\mathbf{r}d\mathbf{r}'\langle \Delta n(\mathbf{r},t)\Delta n(\mathbf{r}',0)\rangle \tag{3.3}$$

The average value of $\Delta n(\mathbf{r},t)$ subject to knowledge of $\Delta n(\mathbf{r}'',0)$ at time 0 can be computed from the phenomenological equations via the Green's function K:

$$\langle \Delta n(\mathbf{r},t)\rangle_0 = \int K(\mathbf{r},\mathbf{r}'',t)\,d\mathbf{r}''\Delta n(\mathbf{r}'',0) \tag{3.4}$$

Under the assumption of charge neutrality, fluctuations in added carrier concentration occur without a change in Coulomb energy. Such fluctuations can then be described in terms of equivalent one body transitions and the methods of Section 12 of Ref. 9 lead to

$$\langle \Delta n(\mathbf{r}',0)\Delta n(\mathbf{r}'',0)\rangle = [\langle (\Delta N)^2\rangle/\Omega]\delta(\mathbf{r}'-\mathbf{r}'') \tag{3.5}$$

where

$$\langle (\Delta N)^2\rangle = NP/(N+P) \tag{3.6}$$

as in Ref. 9, equations (11.11) and (11.17). See also equation (B12), this paper. We may anticipate that even without invoking charge neutrality a careful analysis of correlations in space-charge fluctuations will justify these equations providing all dimensions are large compared to a Debye length (see Appendix C). Equations (3.3), (3.4) and (3.5) can be combined to justify equation (1.9):

$$\Phi(t) = \frac{\langle \Delta N(t)\Delta N(0)\rangle}{\langle (\Delta N)^2\rangle} = \frac{1}{\Omega}\int_\Omega\int_\Omega d\mathbf{r}\,d\mathbf{r}'K(\mathbf{r},\mathbf{r}',t) \tag{3.7}$$

Let us consider first the case in which the volume Ω is part of an infinite homogeneous domain. The Hill–van Vliet experiment[9] in which the noise probes are close together at $x = 0$ and L but far from the ends of the crystal approximates this "fictitious surface" condition. For such an effectively infinite crystal, it is convenient via

$$\Delta n(\mathbf{r},t) = (2\pi)^{-n/2}\int \exp(i\mathbf{k}\cdot\mathbf{r})\Delta n(\mathbf{k},t)\,d\mathbf{k} \tag{3.8}$$

to introduce the Fourier components $\Delta n(\mathbf{k},t)$ because the latter will decay independently:

$$\langle \Delta n(\mathbf{k},t)\rangle = \exp[-\lambda(\mathbf{k})t]\Delta n(k,0) \tag{3.9}$$

If equation (3.9) is inserted into equation (3.8) we arrive at an equation of the form (3.4) and can

deduce that

$$K(\boldsymbol{r}, \boldsymbol{r}', t) = \frac{1}{(2\pi)^n} \int d\boldsymbol{k} \exp[i\boldsymbol{k} \cdot (\boldsymbol{r}-\boldsymbol{r}') - \lambda(\boldsymbol{k})t] \tag{3.10}$$

where $n = 1$, 2 or 3 is the dimensionality of the region Ω.

We may combine equations (1.7), (3.7) and (3.10) to obtain for the noise spectrum

$$g(\omega) = Re \frac{1}{\Omega} \int \frac{d\boldsymbol{k}}{(2\pi)^n} \frac{1}{i\omega+\lambda(\boldsymbol{k})} \left| \int_{\Omega} \exp(i\boldsymbol{k} \cdot \boldsymbol{r})\, d\boldsymbol{r} \right|^2 \tag{3.11}$$

Let us first consider the one-dimensional domain $0 \leqslant x \leqslant L$, then in terms of the new variable $z = kL/2$ equation (3.11) can be written:

$$g(\omega) = Re \frac{1}{\pi} \int_{-\infty}^{\infty} \left(\frac{\sin z}{z}\right)^2 \frac{1}{i\omega+\lambda(z)}\, dz \tag{3.12}$$

where $\lambda(z)$ is simply $\lambda(k)$ rewritten in terms of z.

The phenomenological equation for the case of drift, diffusion and recombination has the form:

$$(\partial/\partial t + \boldsymbol{\Lambda})\Delta n(x, t) = 0 \tag{3.13}$$

with

$$\boldsymbol{\Lambda} = r + v\partial/\partial x - D\partial^2/\partial x^2 \tag{3.14}$$

where r is the recombination rate; v and D are the ambipolar drift velocity and diffusion coefficient respectively. Thus as expected for this "infinite" case, the eigenfunctions of the decay operator $\boldsymbol{\Lambda}$ are plane waves:

$$\boldsymbol{\Lambda} \exp(ikx) = \lambda(k) \exp(ikx) \tag{3.15}$$

where

$$\lambda(k) = r + ivk + Dk^2 \tag{3.16}$$

or

$$\lambda(z) = r + 2i(v/L)z + 4(D/L^2)z^2 \tag{3.17}$$

The integral equation (3.12) has a contribution from the vicinity of $z = 0$ of $r/(\omega^2+r^2)$, precisely the result if $L = \infty$, or if diffusion and drift did not occur. In addition, however, there is a contribution, which for large ω, extends over a large range of z. We can estimate this contribution by setting $\lambda(z) \approx 4(D/L^2)z^2$ and $\sin^2 z \approx \frac{1}{2}$, with the result

$$\Delta g(\omega) \approx Re\left[(2/\pi L) \int_0^{\infty} \frac{dk}{k^2} \frac{1}{i\omega + Dk^2}\right] \tag{3.18}$$

or

$$g(\omega) \approx r/(\omega^2+r^2) + (1/L)(D/2\omega^3)^{1/2} \tag{3.19}$$

in agreement with equation (1.18) since $QS/\Omega = 1/L$. The term in $(1/\omega)^{3/2}$ will dominate the term in $(1/\omega)^2$ when

$$\omega > 2r^2L^2/D = 2r(L/\mathscr{L}_d)^2 \tag{3.20}$$

Since the diffusion length $\mathscr{L}_d = (D/r)^{1/2}$ is usually appreciably smaller than the electrode spacing L the above cross-over frequency can be so high that the boundary diffusion noise is masked by thermal (Nyquist) noise. The ratio of diffusion to thermal noise at this dividing frequency is

$$\frac{G(V,f)_{\text{diff}}}{G(V,f)_{\text{therm}}} = \frac{1}{4}\left(\frac{eV}{kT}\right)^2 \left(\frac{D}{rL^2}\right)^3 \frac{(1+b)^2}{b} \frac{\langle(\Delta N)^2\rangle}{P+Nb} \frac{D_p}{D} \tag{3.21}$$

where D_p is the hole diffusion constant. The inverse three-halves power noise can then best be observed by working with near intrinsic material, and a maximum ratio of $\mathscr{L}_d/L$.

Our approximate evaluation (3.18) of the integral, equation (3.12), can be verified by evaluating the integral

$$J = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2 z}{z^2} dz \frac{Dz^2+2Ez+F}{Az^2+2Bz+C} \tag{3.22}$$

in the complex plane. (See Appendix D.) This integral includes equation (3.12) as a special case with

$$\begin{aligned} &A = 4D/L^2, \quad B = iv/L, \quad C = (i\omega+r) \\ &D = 0, \qquad E = 0, \qquad F = 1 \end{aligned} \tag{3.23}$$

A completely general evaluation is given by equation (D4). However the case in which drift is dominant has already been discussed in detail.[9] We shall therefore discuss here the case of zero drift (which is in any case correct at large ω even in the presence of drift). Using equation (D7) we then find, with no further approximation

$$g(\omega) = ReJ \tag{3.24}$$

$$J = (r+i\omega)^{-1} - [D/(r+i\omega)^3]^{1/2}[1-\exp(-\gamma L)]/L \tag{3.25}$$

where

$$\gamma = [(r+i\omega)/D]^{1/2} \tag{3.26}$$

is the reciprocal ac diffusion length. Equations (1.17) and (1.24) are now verified. Since $Re(\gamma L) \gg 1$ the exponential term which represents an interference effect between the two boundaries may be neglected. We then obtain

$$g(\omega) = \frac{r}{r^2+\omega^2} + \frac{1}{L}\left(\frac{D}{2\omega^3}\right)^{1/2} \times \times \frac{\{[(1+x^2)^{1/2}-x]^{1/2} - x[(1+x^2)^{1/2}+x]^{1/2}\}}{(1+x^2)^{3/2}} \tag{3.27}$$

where $x = r/\omega$. In the region $\omega \gg r$ or $x \ll 1$ equation (3.27) reduces precisely to the result equation (3.19), of our more heuristic treatment of equation (3.18). We may feel justified, then, in applying this treatment in two and three dimensions to verify equation (1.18) for the asymptotic behavior of the noise.

For a sphere, with $\Omega = 4\pi R^3/3$

$$\int_\Omega \exp(i\boldsymbol{k}\cdot\boldsymbol{r})\, d\boldsymbol{r} = \frac{4\pi}{k^3}(\sin kR - kR\cos kR) \tag{3.28}$$

and keeping only the dominant terms (with $\cos^2 kR \approx \frac{1}{2}$)

$$\left|\int_\Omega \exp(i\boldsymbol{k}\cdot\boldsymbol{r})\, d\boldsymbol{r}\right|^2 \approx \frac{8\pi^2 R^2}{k^4} \tag{3.29}$$

Setting $\lambda(\boldsymbol{k}) \approx Dk^2$ in equation (3.11) we get an integral identical with equation (3.18) if $2/L$ is replaced by $3/R$ thus verifying equations (1.18) and (1.12) for a sphere. The result, equation (1.14), for a circle follows from a similar treatment based on the asymptotic properties of Bessel functions.

For the case of a rectangle, one must first succeed in deriving the lemma:

$$\int_0^{2\pi} \frac{\sin^2(\frac{1}{2}kL_1\cos\theta)\sin^2(\frac{1}{2}kL_2\sin\theta)}{\cos^2\theta\sin^2\theta} \times \times d\theta \to k(L_1+L_2)(\pi/2) \tag{3.30}$$

for large k, and then our results, equations (1.18) and (1.13), follow immediately from an integral of the form equation (3.18).

4. THE INFLUENCE OF BOUNDARY CONDITIONS

To what extent does the $(1/\omega)^{3/2}$ law found to be valid for "fictitious" boundaries of a variety of shapes remain valid for diffusion effects at actual boundaries of solids?

A first example might be the boundary condition that the normal gradient of the carrier density vanishes at the surface. In this case diffusion across the surface is prevented, and the spectrum will simply be that associated with volume recombination $r/(r^2+\omega^2)$. (The lowest mode of decay of the system will be spatially uniform, with decay constant r, all other modes will be orthogonal and hence will not affect fluctuations in the total number of particles.)

A second example might be a filament, with zero gradient at the sides of the filament, noise probes at $x = 0$ and L, but voltage contacts far removed from the noise probes. In this case, only the modes with densities uniform in the y- and z-direction will contribute, and the result is the same as if the specimen had been infinite in the y- and z-directions i.e. it is equivalent to the one-dimensional problem solved in Section 3 and has the expected $(1/\omega)^{3/2}$ behavior at high frequencies because of diffusion across the boundaries at $x = 0$ and L.

As a third example, let us consider surface recombination, in a one-dimensional case, with boundaries at $x = \pm B$. Our Green's function as usual (cf. Ref. 9, equation (11.20)) obeys

$$(\partial/\partial t + \Lambda)K(x, x', t) = \delta(x-x')\delta(t) \tag{4.1}$$

with

$$\Lambda = r - D\partial^2/\partial x^2 \tag{4.2}$$

since we are concerned primarily with the effects of diffusion. In addition K obeys the boundary conditions

$$\begin{aligned} D\partial K/\partial x &= sK \text{ at } x = -B \\ D\partial K/\partial x &= -s'K \text{ at } x = B \end{aligned} \tag{4.3}$$

appropriate to surface recombination velocities s and s'. In view of equations (1.7) and (1.9) the spectral distribution can be expressed in the form

$$g(\omega) = Re\frac{1}{2B}\int_{-B}^{B}\int_{-B}^{B} dx\, dx' g(x, x', \omega) \tag{4.4}$$

where $g(x, x', \omega)$ is the Laplace transform

$$g(x, x', \omega) = \int_0^\infty \exp(-i\omega t)K(x, x', t)\, dt \qquad (4.5)$$

and consequently obeys

$$(\Lambda + i\omega)g(x, x', \omega) = \delta(x-x') \qquad (4.6)$$

or

$$(d^2/dx^2 - \gamma^2)g(x, x', \omega) = -(1/D)\delta(x-x') \qquad (4.7)$$

where

$$\gamma^2 = (r+i\omega)/D \qquad (4.8)$$

Comparison of equations (4.1) and (4.6) indicates that $g(x, x', \omega)$ may be interpreted as the response at x to a single frequency $\exp(i\omega t)$ source at x'.

Integration of equation (4.7) yields

$$\int_{-B}^{B} g(x, x')\, dx = (1/\gamma^2 D) + \\ + (1/\gamma^2)[g_x(B, x') - g_x(-B, x')] \qquad (4.9)$$

where $g_x(x, x') \equiv \partial g(x, x')/\partial x$ and the first term is the volume term which would remain present if the second derivative diffusion terms were discarded. For any interior point x', $g_x(\pm B, x')$ can be regarded as a source free solution of an equation like (4.7) on the variable x'. Thus the second integration in equation (4.4) can be explicitly carried out:

$$g(\omega) = Re\{(1/\gamma^2 D) + (2B\gamma^4)^{-1}[g_{xx'}(B, B) + \\ + g_{xx'}(-B, -B) - g_{xx'}(B, -B) - g_{xx'}(-B, B)]\} \qquad (4.10)$$

where

$$g_{xx'}(x, x') \equiv \partial^2 g(x, x')/\partial x \partial x' \qquad (4.11)$$

If the differential equation (4.7) and boundary conditions (4.3) possess inversion symmetry about $x = 0$, we can simplify equation (4.10) to:

$$g(\omega) = r/(r^2+\omega^2) + \\ + Re\{(B\gamma^4)^{-1}[g_{xx'}(B, B) - g_{xx'}(B, -B)]\} \qquad (4.12)$$

Equations (4.10) and (4.12) are the basis for our intuitive remarks in the introduction that the noise automatically splits into a volume term plus a correction for diffusion across or into the boundary. Furthermore, equation (4.10) explicitly displays a term for diffusion across each boundary plus two interference terms which are of negligible importance when $\gamma B \gg 1$, i.e. when the boundaries are far apart compared to an ac diffusion length.

The Green's function solution of equation (4.7) can be found by the standard method[15] of setting

$$\begin{aligned} g(x, x') &= Cu(x)v(x') \qquad x < x' \\ &= Cu(x')v(x) \qquad x > x' \end{aligned} \qquad (4.13)$$

requiring $u(x)$ and $v(x)$ to be solutions of the homogeneous equation obeying the left- and right-hand boundary conditions respectively, and choosing C so that the jump in $\partial g/\partial x$ at x':

$$[\partial g/\partial x]_{x'-}^{x'+} = -1/D \qquad (4.14)$$

agrees with that required by the source. The result for surface recombination velocity boundary conditions equation (4.3), is:*

$$\begin{aligned} u(x) &= \cosh\gamma x + \alpha \sinh\gamma x \\ v(x) &= \cosh\gamma x - \alpha' \sinh\gamma x \end{aligned} \qquad (4.15)$$

$$C = [\gamma D(\alpha+\alpha')]^{-1} \qquad (4.16)$$

where

$$\alpha = \frac{\sinh\gamma B + y\cosh\gamma B}{\cosh\gamma B + y\sinh\gamma B}; \qquad \alpha' = \frac{\sinh\gamma B + y'\cosh\gamma B}{\cosh\gamma B + y'\sinh\gamma B} \qquad (4.17)$$

and

$$y = (s/\gamma D); \qquad y' = (s'/\gamma D) \qquad (4.18)$$

The second term in equation (4.10) for the surface contribution to the noise can now be written exactly as in equation (1.17):

$$\Delta g(\omega) = -[2/(2B)]Re\{[D/(r+i\omega)^3]^{1/2}Q\} \qquad (4.19)$$

where Q is given without approximation by

$$Q = \frac{2yy' + (y+y')\coth(\gamma B)}{2(1+yy')\coth(\gamma B) + (y+y')[1+\coth^2(\gamma B)]} \qquad (4.20)$$

* The authors are indebted to H. K. GUMMEL for independently verifying these results for the case $s = s'$ and for calling our attention to CHAMPLIN's work[12] on the same subject.

When $y' = y$, this reduces to

$$Q = y/(1+y\coth\gamma B) \xrightarrow[B\to\infty]{} y/(1+y) \qquad (4.21)$$

in agreement with equation (1.25) and CHAMPLIN[12]. We suggest that the limiting result $Q = y/(1+y)$ be regarded as the probability that an electron which hits the surface does not return. In support of this idea, we note that when $\gamma B \gg 1$, equation (4.20) reduces to

$$Q = \tfrac{1}{2}[y/(1+y)]+\tfrac{1}{2}[y'/(1+y')] \qquad (4.22)$$

where the factor $\frac{1}{2}$ arises because we have used $S/\Omega = 2/(2B)$ in equation (4.19) whereas we should have used $1/(2B)$ for each boundary. Indeed, the first term in equation (4.22) arises from $g_{xx}(-B, -B)$, and the second from $g_{xx'}(B, B)$ in the limit $\gamma B \gg 1$. The interference terms are of order $\exp(-2\gamma B) \ll 1$. Further support for our interpretation can be garnered from a calculation of the semi-infinite case using $v(x) = \exp(-\gamma x)$, $C = [\gamma D(1+\alpha)]^{-1}$ which leads directly to the probability $y/(1+y)$.

(We also note that the appropriate Green's function for the infinite domain (vanishing as $x \to \pm\infty$) is:

$$g(x, x') = (2\gamma D)^{-1}\exp(-\gamma|x-x'|) \qquad (4.23)$$

when inserted into equation (4.10) or (4.12) leads to complete agreement with the result, equation (3.25) for the "fictitious" boundary calculated by an entirely different method.)

The surface recombination result, equation (4.21), leads to a dominant inverse three-halves behavior

$$g(\omega) \sim (D/2\omega^3)^{1/2}/B \qquad (4.24)$$

in the region

$$8r^2B^2/D < \omega < s^2/2D \qquad (4.25)$$

whereas at higher frequencies

$$g(\omega) \approx [r+(s/B)]/\omega^2 \qquad (4.26)$$

corresponding to an enhanced volume recombination. The $(1/\omega)^{3/2}$ region can therefore only be observed if the recombination velocity is large enough:

$$s > 4rB \qquad (4.27)$$

A more detailed understanding of the shift from a $(1/\omega)^{3/2}$ to a $(1/\omega)^2$ behavior can be obtained if one expands by standard methods[16] the Green's function $K(x, x', t)$ or $g(x, x', \omega)$ for the case $s = s'$ in terms of the eigenfunctions $\cos k_n x$ of $\mathbf{\Lambda}$, where the k_n obey the boundary condition

$$\tan k_nB = (s/k_nD) \qquad (4.28)$$

The final result for $g(\omega)$ takes the form:

$$g(\omega) = Re\frac{2}{B^2}\sum_{n=1}^{\infty}\frac{\sin^2(k_nB)}{k_n^2}\times \times\frac{[1-\sin(2k_nB)/(2k_nB)]^{-1}}{i\omega+r+Dk_n^2} \qquad (4.29)$$

The behavior of $g(\omega)$ for large ω is controlled by the high modes $Dk_n^2 \sim \omega$. The factor in brackets is a minor correction and can be omitted. If one replaced the sum by an integral using $\Delta k \approx \pi/B$, the integral would be identical to equation (3.12) and reduce to the fictitious boundary result. This replacement is not valid because $\sin^2(k_nB)$ does not sample $\sin^2(kB)$ at closely spaced points but only at points roughly half a wavelength apart. For this same reason $\sin^2 k_nB$, which represents the squared density gradient $(\partial u_n/\partial x)^2$ at the boundary, cannot be replaced by the volume average value of $\frac{1}{2}$ appropriate to the fictitious boundary case. Rather one must use equation (4.28) to obtain

$$\sin^2 k_nB = [1+(k_nD/s)^2]^{-1} \qquad (4.30)$$

from which we see that the mean square density gradient at the boundary is about unity for the low modes whereas it decreases for the high modes. In other words, for the low modes one has "infinite" surface recombination and the surface is an antinode for the density gradient leading to a mean squared density of twice the volume average and a sticking probability Q of unity, twice the result for a fictitious boundary. On the other hand, equation (4.29) indicates that for any fixed s, the surface must approach becoming a node for the density gradient for sufficiently large n. In other words the surface recombination velocity boundary condition inhibits diffusion into the boundary at very high frequencies or propagation constants.

This last result indicates that a surface recombination velocity boundary condition is not entirely realistic unless one allows s to become frequency dependent. For example, let us suppose

that recombination occurs not at a surface but in a region of width a with recombination coefficient r' and diffusion coefficient D'. We may suppose that $x = 0$ is an ideal surface with $\partial n/\partial x = 0$. The boundary condition at $x = a$ in contact with the interior is then

$$s/D = (\partial n/\partial x)/n = \gamma' \tanh(\gamma' a) \tag{4.31}$$

where

$$\gamma' = [(r'+i\omega)/D']^{1/2} \tag{4.32}$$

and we have used $n(x) = \cosh(\gamma' x)$ in the "surface region". Thus

$$s \approx (D/D')(r'+i\omega)a \sim (D/D')r'a \text{ for } \gamma' a < 1 \tag{4.33}$$

$$s \approx [(r'+i\omega)/D']^{1/2} D \text{ for } \gamma' a > 1 \tag{4.34}$$

Although equation (4.33) may be valid for a broad frequency range, the limiting behavior at very high frequencies (or at all frequencies when $a = \infty$) is determined by (4.34):

$$y = (s/\gamma D) \approx \gamma'/\gamma \to (D/D')^{1/2} \tag{4.35}$$

$$Q = y/(y+1) \approx \gamma'/(\gamma+\gamma') \to \sqrt{D}/(\sqrt{D}+\sqrt{D'}) \tag{4.36}$$

so that the probability Q of no return diffusion is unity when $D' = 0$ and $\frac{1}{2}$ when $D' = D$, in agreement with our fictitious boundary result. In any case Q becomes independent of frequency at high frequencies and a $(1/\omega)^{3/2}$ law prevails at high frequencies whenever diffusion across the boundary is a possible process.

5. TRAPPING PLUS DIFFUSION

A general conservation equation in the presence of trapping, recombination, drift and diffusion has been derived by VAN ROOSBROECK[8], assuming only charge neutrality. In its general, nonlinear form, this equation can be written as:

$$\partial m/\partial t = \operatorname{div}(\hat{D}_n \operatorname{grad} n + \hat{D}_p \operatorname{grad} p) - \\ - \hat{\nu}_n \cdot \operatorname{grad} n - \hat{\nu}_p \cdot \operatorname{grad} p + G - R \tag{5.1}$$

where

$$m \equiv n + \hat{n} = p + \hat{p} \tag{5.2}$$

$$\hat{D}_n = pD_pD_n/(nD_n+pD_p) = Dp/(n+p) \tag{5.3}$$

$$\hat{D}_p = nD_pD_n/(nD_n+pD_p) = Dn/(n+p) \tag{5.4}$$

$$D = (n+p)D_pD_n/(nD_n+pD_p) = \hat{D}_n+\hat{D}_p \tag{5.5}$$

$$\hat{\nu}_n = -(e\mu_p\mu_n p/\sigma^2)\boldsymbol{I} = p\nu/(p-n) \tag{5.6}$$

$$\hat{\nu}_p = (e\mu_p\mu_n n/\sigma^2)\boldsymbol{I} = -n\nu/(p-n) \tag{5.7}$$

$$\nu = e\mu_n\mu_p(n-p)\sigma^{-2}\boldsymbol{I} = \nu_n+\nu_p \tag{5.8}$$

Here D_n and D_p are the concentration independent electron and hole diffusion constants, D and ν are the concentration dependent ambipolar diffusivity and drift velocity respectively, $\hat{n}$ and $\hat{p}$ are the concentrations of trapped electrons and holes, and G and R are the rates of generation and recombination of electron hole pairs. The rate of change of free electrons can be determined from

$$\partial n/\partial t = \partial m/\partial t - \partial \hat{n}/\partial t \tag{5.9}$$

where the rate of change of trapped carriers is determined by the usual equation (2.11).

Equation (5.1) can be linearized by making the replacement

$$n \to n+\Delta n, \qquad p \to p+\Delta p \tag{5.10}$$

where n and p now represent (possibly space dependent) *steady* (i.e. time independent) solutions in the absence of fluctuations or variable driving forces—although it would not change our analysis if these steady values were modified by the presence of a steady light, or the presence of shallow donors and acceptors. We shall for the purposes of this paper explicitly assume homogeneity for the steady solutions in the form

$$\operatorname{grad} n = \operatorname{grad} p = 0 \tag{5.11}$$

In this case, the linearization of equation (5.1) is accomplished trivially by replacing n and p underneath the gradients by Δn and Δp, with $\hat{D}_n$ and $\hat{D}_p$ taking the now concentration independent values appropriate to the steady-state. In addition, we shall assume either that no trapped holes are present, or that their fluctuations may be neglected: $\Delta\hat{p} = 0$, and impose charge neutrality in the form:

$$\Delta p = \Delta n + \Delta\hat{n} \tag{5.12}$$

Furthermore, we shall assume that our material is sufficiently p-type for us to approximate the net recombination by

$$R - G \approx r\Delta n \tag{5.13}$$

although there would in fact be no significant increase in labor to keep a term in Δp.

Making use of equation (2.13), (5.1), (5.9), (5.12) and (5.13) we obtain our final linearized equations in the form

$$\begin{aligned}\partial\Delta n/\partial t &= -(\nu+r+\mathbf{v}\cdot\nabla-D\nabla^2)\Delta n+\\ &\quad+(g'-\dot{\mathbf{v}}_p\cdot\nabla+\hat{D}_p\nabla^2)\Delta\hat{n}\\ \partial\Delta\hat{n}/\partial t &= \nu\Delta n-g'\Delta\hat{n}\end{aligned}\tag{5.14}$$

where

$$\nu=(\hat{n}_t-\hat{n})C$$

is the capture rate into the empty traps, as in equation (2.16),

$$g'=g+Cn\tag{5.15}$$

is the effective release rate, as in equation (2.18), and C represents a capture rate per empty trap (cross section x velocity) with dimensions cm^3/sec.

Since the relation between solutions in one, two and three dimensions has been found to be a fairly simple one (cf. equation 1.17), and we understand from Section 4 the effect of modifying the boundary conditions, we can best isolate the effects of trapping by considering our usual one-dimensional example with "fictitious" boundaries at $x=0$ and L.

One can proceed easily by expanding in terms of eigenfunctions of the decay operator. For our effectively infinite crystal, these eigenfunctions are plane waves, and as in equation (3.8) we introduce the Fourier coefficients $\Delta n(k,t)$, $\Delta\hat{n}(k,t)$ which in view of equation (5.14) obey the equations:

$$\begin{aligned}\partial\Delta n(k,t)/\partial t &= -(\nu+r+i\nu k+Dk^2)\Delta n(k)+\\ &\quad+(g'-i\hat{v}_p k-\hat{D}_p k^2)\Delta\hat{n}(k)\\ \partial\Delta\hat{n}(k,t)/\partial t &= \nu\Delta n(k)-g'\Delta\hat{n}(k)\end{aligned}\tag{5.16}$$

A method for handling these equations exactly and with no computational difficulties is outlined in Appendix E. However, we shall gain more physical insight, and obtain simpler formulae (which can only be obtained with appreciable difficulty from the exact results) by making approximations based on the method of low mode isolation developed in Section 2.

Following the procedure of equation (2.31), the low mode isolation is accomplished by setting $\partial\Delta n/\partial t=0$ or

$$\Delta n(k)=\Delta\hat{n}(k)(g'-i\hat{v}_p k-\hat{D}_p k^2)/(\nu+r+i\nu k+Dk^2)\tag{5.17}$$

so that the trapped carriers experience the slow decay:

$$\partial\Delta\hat{n}(k,t)/\partial t=-\lambda_S(k)\Delta\hat{n}(k,t)\tag{5.18}$$

with

$$\lambda_S(k)=\frac{g'r+i\nu v_B k+Dv_A k^2}{\nu+r+i\nu k+Dk^2}\tag{5.19}$$

$$v_A=g'+\nu n/(n+p)\tag{5.20}$$

$$v_B=g'-\nu n/(p-n)\tag{5.21}$$

which reduces for $k=0$ to equation (2.33).

On the other hand, the fast decay, which essentially characterizes the free electron decay can be obtained from the trace as in equation (2.27):

$$\lambda_F(k)\approx\nu+r+g'+i\nu k+Dk^2\tag{5.22}$$

From equation (3.20) we see that the effects of diffusion are not important unless

$$\omega>2\nu^2L^2/D\sim 6\times10^{11}/\text{sec}\tag{5.23}$$

i.e. as claimed in Section 1, diffusion (and drift) can be neglected entirely with regard to the fast modes.

(Returning to equation (5.19) for the slow modes, we *might* be tempted to reduce our problem to an equivalent problem with no traps by setting $k=0$ in the denominator, neglecting terms like Dk^2 with respect to ν, however this approximation lets $\lambda_S(k)\to\infty$ as $k\to\infty$, whereas in fact it approaches v_A. This approximation will therefore be valid *only* for $\omega<v_A$.)

In view of remark (6) in the introduction and equation (2.10), we have, keeping only the $\hat{N}$ contributions:

$$G(V,f)=4[\hat{N}'V^2/(P+Nb)^2]Re\,g(\omega)\tag{5.24}$$

when

$$g(\omega)=\int_0^\infty\exp(-i\omega t)\hat{\Phi}(t)\,dt\tag{5.25}$$

$$\hat{\Phi}(t)=\frac{1}{L}\int_0^L\int_0^L\hat{K}(x,x',t)\,dx\,dx'\tag{5.26}$$

VOL. 14 1960

where $\hat{K}$, the Green's function for the trapped carriers, can be defined as in equation (3.10) using $\lambda_S(k)$ as the decay constant. Our result can then be brought into the form of equation (3.22) with the parameters given by

$$\begin{aligned} A &= 4(D/L^2)(\nu_A+i\omega); & D &= 4(D/L^2) \\ B &= i(v/L)(\nu_B+i\omega); & E &= i(v/L) \\ C &= (\nu+r)(\lambda_S+i\omega); & F &= \nu+r \end{aligned} \tag{5.27}$$

where $\lambda_S \approx rg'/(\nu+r)$ is the slow decay rate in the absence of drift and diffusion.

We first discuss the case of negligible drift, using equation (D7). As long as L is significantly larger than a diffusion length, the exponential interference term in equation (D7) can be dropped and we have:

$$g(\omega) = \frac{\lambda_S}{(\lambda_S)^2+\omega^2} - \\ -\frac{1}{L}\left(\frac{D}{\nu+r}\right)^{1/2} Re\frac{\nu_A-\lambda_S}{[(\nu_A+i\omega)(\lambda_S+i\omega)^3]^{1/2}} \tag{5.28}$$

The approximation in equation (5.19) of omitting Dk^2 in the denominator would have yielded equation (5.28) with $\nu_A-\lambda_S$ replaced by ν_A, and $\nu_A+i\omega$ replaced by ν_A, the latter valid as long as $\omega \ll \nu_A$. This approximation suggests an effective diffusion coefficient

$$D_{\text{eff}} = D\nu_A/(\nu+r) \tag{5.29}$$

and indeed in the region $\lambda_S \ll \omega \ll \nu_A$, equation (5.4) reduces to the familiar form

$$\Delta g(\omega) = (1/L)(D_{\text{eff}}/2\omega^3)^{1/2} \tag{5.30}$$

for the diffusion contribution to the noise. In fact λ_S plays the role of an effective recombination coefficient, and over the range $0 \leqslant \omega \ll \nu_A$ we have

$$\Delta g(\omega) = \frac{1}{L}\left(\frac{D_{\text{eff}}}{2\omega^3}\right)^{1/2} \times \\ \times \frac{\{[(1+x^2)^{1/2}-x]^{1/2}-x[(1+x^2)^{1/2}+x]^{1/2}\}}{(1+x^2)^{3/2}} \tag{5.31}$$

where $x = \lambda_S/\omega$, in agreement with equation (3.27).

An approximation valid in the overlapping range $\lambda_S \ll \omega$ is

$$\Delta g(\omega) = \frac{1}{L}\left(\frac{D_{\text{eff}}}{2\omega^3}\right)^{1/2}\left[\frac{(1+y^2)^{1/2}+y}{1+y^2}\right]^{1/2} \tag{5.32}$$

where $y = \omega/\nu_A$. In the common region $\lambda_S \ll \omega \ll \nu_A$, equations (5.32) and (5.31) reduce to the simple inverse 3/2 law. (See Fig. 2.) The ratio of the maximum frequency ν_A at which the 3/2 law occurs to the minimum frequency defined by equation (3.20) at which it dominates is

$$\frac{\nu_A}{2\lambda_S^2 L^2/D_{\text{eff}}} \approx \left(\frac{g'+(n/p)\nu}{g'}\right)^2 \frac{D\nu}{2L^2r^2} \approx 200 \tag{5.33}$$

for the choice of parameters in Table 1, although it could be significantly larger for different choices of parameters. In Fig. 2 the slope appears to be close to 3/2 over about three decades.

In the region $\omega \gg \nu_A$, equation (5.32) reduces to an inverse square law

$$\Delta g(\omega) \approx (D_{\text{eff}}\nu_A)^{1/2}/(L\omega^2) \tag{5.34}$$

which is greater than the high frequency contribution (second term of equation 1.4) up to about

$$\omega \sim [(D_{\text{eff}}\nu_A)^{1/2}\nu\bar{N}^r/(NL)]^{1/2} \sim 10^6 \tag{5.35}$$

as may be seen in Fig. 3. It is, of course, not necessary to neglect the fast decay contribution. One need only add the $\Delta g(\omega)$ obtained here for the diffusion correction of the slow modes to the exact volume contribution, equation (2.22) of both slow and fast modes.

Let us now discuss the situation in which drift is the dominating influence by neglecting diffusion. Using equation (D9) the noise $g(\omega)$ is the real part of

$$J = \frac{1}{\lambda_S+i\omega} \mp \frac{v}{L}\frac{\nu_B-\lambda_S}{(\nu+r)(\lambda_S+i\omega)^2} \times \\ \times\left\{1-\exp\left[\mp\frac{(\nu+r)(\lambda_S+i\omega)}{(v/L)(\nu_B+i\omega)}\right]\right\} \tag{5.36}$$

where the sign is chosen so that the exponential has magnitude less than unity. When the field is low enough for the interference term to be neglected,

equation (5.36) reduces for $\omega \gg \lambda_S$ to

$$g(\omega) = ReJ \approx [\lambda_S + (v/L)(\nu_B - \lambda_S)/(\nu + r)]/\omega^2 \tag{5.37}$$

In the limit as the field approaches infinity, B and E approach infinity with a finite ratio. Equation (D10) then yields

$$g(\omega) = ReJ \approx Re(E/B) = Re(\nu_B + i\omega)^{-1} \tag{5.38}$$

a simple relaxation spectrum.

This result is unusual in that ν_B as defined by equation (5.21) can go negative, i.e. one has *amplification* rather than decay. We can, however, using $g' = g + Cn$, $C = \nu/(\hat{n}_t - \hat{n})$, $g = \nu n/\hat{n}$, (cf. equations 5.14, 5.15 and 2.14) write:

$$\nu_B = \nu n[1/\hat{n} + 1/(\hat{n}_t - \hat{n}) - 1/(p - n)] \tag{5.39}$$

so that ν_B can be made negative only if the number of *majority carriers* (more precisely p–n) can be made smaller than the number of filled and the number of empty traps. This is possible in an intrinsic region.

The limiting result, equation (5.38), can be obtained from equation (5.36) by expanding the exponential and keeping only the first nonvanishing contribution. Thus the limiting result (equation 5.38) and the possibility of amplification require a field large enough so that

$$v/L > (\nu + r)[\lambda_S^2 + \omega^2)/(\nu_B^2 + \omega^2)]^{1/2} \tag{5.40}$$

or

$$v/L > (\nu + r)\lambda_S/\nu_B \sim r, \qquad \omega < \lambda_S \tag{5.41}$$

$$v/L > (\nu + r)\omega/\nu_B, \qquad \lambda_S < \omega < \nu_B \tag{5.42}$$

$$v/L > \nu + r, \qquad \nu_B < \omega \tag{5.43}$$

Equation (5.41) using Table 1 requires modest fields of the order of 1 V/cm, but equation (5.43) requires fields of the order of 1000 V/cm.

Acknowledgements—The authors would like to acknowledge discussions on Debye-Hückel theory (Appendix C) with Dr. P. A. WOLFF of these Laboratories who has made a quantum-mechanical calculation of the correlations by summation of ring diagrams[18] which reduces for separations large compared to a thermal de Broglie wavelength to the classical results equations (C1)–(C5).

APPENDIX A

Kinetic derivation of concentration fluctuations

We adopt as our basic transition probability per unit time

$$\begin{aligned} w(N', \hat{N}'; N, \hat{N}) = {} & [G\delta(N', N+1) + \\ & + R\delta(N', N-1)]\delta(\hat{N}', \hat{N}) + \\ & + g\hat{N}\delta(N', N+1)\delta(\hat{N}', \hat{N}-1) + \\ & + (C/\Omega)(N_t - \hat{N})N\delta(N', N-1)\delta(\hat{N}', \hat{N}+1) \end{aligned} \tag{A1}$$

where N', $\hat{N}'$ are the final number of free and bound carriers. Then, summing over the primed variables:

$$\begin{aligned} \sum (N' - N)w(N', \hat{N}'; N, \hat{N}) &= G - R + g\hat{N} - \nu N \\ \sum (\hat{N}' - \hat{N})w(N', \hat{N}'; N, \hat{N}) &= \nu N - g\hat{N} \end{aligned} \tag{A2}$$

where

$$\nu = (N_t - \hat{N})C/\Omega \tag{A3}$$

yields the right-hand side of equation (2.11) in accord with Ref. 9, equation (5.22). According to Ref. 9, equation (5.23) the elements of the diffusion matrix are defined by

$$\begin{aligned} D_{11} &= \tfrac{1}{2}\sum (N' - N)^2 w; \\ D_{22} &= \tfrac{1}{2}\sum (\hat{N}' - \hat{N})^2 w; \\ D_{12} &= \tfrac{1}{2}\sum (\hat{N}' - \hat{N})(N' - N)w \end{aligned} \tag{A4}$$

e.g.

$$D_{11} = \tfrac{1}{2}[(G + R) + (g\hat{N} + \nu N)] \tag{A5}$$

$$= R + g\hat{N} \approx rN + g\hat{N} \tag{A6}$$

where we evaluate D_{11} at the steady-state value and use the steady-state conditions

$$G = R, \qquad \nu N = g\hat{N} \tag{A7}$$

to obtain equation (A6). We may take the "total" recombination rate per electron R/N as approximately equal to the (observed) differential rate r defined by $R - G = r\Delta N$ when $N \ll P$. The complete matrix, obtained in this way is

$$\boldsymbol{D} = \begin{pmatrix} Nr + \hat{N}g, & -\hat{N}g \\ -\hat{N}g, & \hat{N}g \end{pmatrix} \tag{A8}$$

and

$$\boldsymbol{\Lambda}^{-1}\boldsymbol{D} = \begin{pmatrix} N, & 0 \\ (\nu N - \hat{N}g)/g', & \hat{N}g/g' \end{pmatrix} \tag{A9}$$

In view of equations (A7) and (A3), the 21 element of this matrix vanishes and

$$g/g' = 1 - (\hat{N}/N_t) \tag{A10}$$

so that

$$\mathbf{\Lambda}^{-1}\mathbf{D} = \begin{pmatrix} N, & 0 \\ 0, & \hat{N}' \end{pmatrix}$$

Since $\mathbf{\Lambda}^{-1}\mathbf{D}$ is symmetric, our results are in accord with an underlying time reversibility,* and the second moments are by Ref. 9, equation 5.20 given by $\mathbf{\Lambda}^{-1}\mathbf{D}$ itself, thus justifying the moments quoted in equation (2.20).

APPENDIX B

Fluctuations in occupation number with finite hole reservoir

We shall show that the results of Appendix A, derived without explicit consideration of holes remain valid as long as the number of holes is large compared to the number of traps and the number of free carriers. For this purpose, we apply equation (12.38) of Ref. 9 using Fermi statistics in the form:

$$\langle \Delta n(a)\Delta n(c)\rangle = n(c)'\delta(a,c) - [n(a)'n(c)'/\textstyle\sum_b n(b)'] \tag{B1}$$

where

$$\begin{aligned} n(a)' &= (\text{occupancy}) \times (\text{fraction empty}) \\ &= n(a)\{1-[n(a)/n(a)_{\max}]\} \end{aligned} \tag{B2}$$

Equation (B1) has been constructed to describe the situation in which the occupancies are assumed to be subject to the constraint of a fixed total number. Let us apply this result assuming there are three states $n(1) = N =$ number of conduction electrons, $n(2) = \hat{N} =$ number of trapped electrons, $n(3) =$ number of electrons in the valence band $= N_v - P$ when $N_v =$ number of valence states and $P =$ number of holes. Thus

$$n(1)' = N[1-N/N_c] \approx N \tag{B3}$$

$$n(2)' = \hat{N}[1-\hat{N}/N_t] = \hat{N}' \tag{B4}$$

$$n(3)' = (N_v-P)(P/N_v) \approx P \tag{B5}$$

assuming nondegeneracy for both electrons and holes. Remembering that $\Delta n(3) = -\Delta P$, we can now write our results in the form:

$$\langle(\Delta N)^2\rangle = N - N^2/(P+N+\hat{N}') \tag{B6}$$

$$\langle(\Delta\hat{N})^2\rangle = \hat{N}' - (\hat{N}')^2/(P+N+\hat{N}') \tag{B7}$$

$$\langle\Delta N\Delta\hat{N}\rangle = -N\hat{N}'/(P+N+\hat{N}) \tag{B8}$$

$$\langle(\Delta P)^2\rangle = P(\hat{N}'+N)/(P+N+\hat{N}') \tag{B9}$$

$$\langle\Delta P\Delta N\rangle = NP/(P+N+\hat{N}') \tag{B10}$$

$$\langle\Delta P\Delta N\rangle = \hat{N}'P/(P+N+\hat{N}') \tag{B11}$$

* That the symmetry of $\mathbf{\Lambda}^{-1}\mathbf{D}$ is sufficient (as well as necessary) for time reversibility may be seen by combining equations (5.19) and (6.4) of Ref. 9.

These results are clearly consistent with the constraint $\Delta P = \Delta N + \Delta\hat{N}$ on a fixed total number of electrons and they reduce to previous results when $P \gg N$ and $P \gg \hat{N}'$. They also reduce to

$$\langle(\Delta N)^2\rangle = \langle(\Delta P)^2\rangle = NP/(N+P) \tag{B12}$$

when no traps are present. When $P = 0$, the results agree with Ref. 9, equations (12.41) and (4.30) for fluctuations in majority carriers supplied by traps.

APPENDIX C

Correlations introduced by coulomb forces

For the concentrations appropriate to nondegenerate electrons and holes in semiconductors, it is adequate to use the classical Debye–Hückel[17,18] theory which yields

$$\langle\Delta n(\mathbf{r})\Delta n(\mathbf{r}')\rangle = n(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') - [n^2/(n+p)]F(\mathbf{r}-\mathbf{r}') \tag{C1}$$

$$\langle\Delta p(\mathbf{r})\Delta p(\mathbf{r}')\rangle = p(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}') - [p^2/(n+p)]F(\mathbf{r}-\mathbf{r}') \tag{C2}$$

$$\langle\Delta n(\mathbf{r})\Delta p(\mathbf{r}')\rangle = [np/(n+p)]F(\mathbf{r}-\mathbf{r}') \tag{C3}$$

where

$$F(\mathbf{r}) = \lambda^2 \exp(-\lambda r)/(4\pi r) \tag{C4}$$

and

$$\lambda^2 = 4\pi(n+p)e^2/(kT) \tag{C5}$$

defines the extrinsic Debye screening length $1/\lambda$. As long as the above correlations are integrated against a function which is smooth over distances large compared to an extrinsic Debye length, it is adequate to make the replacement

$$F(\mathbf{r}-\mathbf{r}') \approx \delta(\mathbf{r}-\mathbf{r}') \tag{C6}$$

which yields the result of equation (3.5) computed under the assumption of charge neutrality:

$$\langle\Delta n(\mathbf{r})\Delta n(\mathbf{r}')\rangle \approx [np/(n+p)]\delta(\mathbf{r}-\mathbf{r}') \tag{C7}$$

The error involved in equation (C7) can be most readily appreciated by rewriting equation (C1) in terms of the fluctuations in the Fourier component (3.8):

$$\begin{aligned} &\langle\Delta n(\mathbf{k})\Delta n(\mathbf{k}')\rangle \\ &= \delta(\mathbf{k}-\mathbf{k}')\{np/(n+p) - [n^2/(n+p)]k^2/(\lambda^2+k^2)\} \end{aligned} \tag{C8}$$

Thus the correction term is only important for the high Fourier components $k \sim \lambda$, i.e. at frequencies

$$\begin{aligned} \omega \sim D\lambda^2 &= D(4\pi e^2/kT)(n+p) \\ &= 4\pi(n+p)e\mu \approx 4\pi\sigma \end{aligned} \tag{C9}$$

comparable to the reciprocal dielectric relaxation time.

(In the presence of trapping—cf. equation (1.29)—the effective diffusion constant is smaller by a factor ν_A/ν, as is the frequency at which charge fluctuations make their effects felt.)

APPENDIX D

An integral

If $F(z)$ has no singularities other than simple poles and is bounded at infinity then

$$J = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin^2 z}{z^2}\, dz F(z) \tag{D1}$$

$$= F(0) + \tfrac{1}{2}i\{\textstyle\sum_a R(z_a)[1-\exp(2iz_a)]/z_a^2 - \\ - \textstyle\sum_b R(z_b)[1-\exp(-2iz_b)]/z_b^2\} \tag{D2}$$

where z_a, (z_b) are the poles of $F(z)$ in the upper (lower) half-plane and $R(z_a)$, ($R(z_b)$) are the corresponding residues at these points. If

$$F(z) = (Dz^2+2Ez+F)/(Az^2+2Bz+C) \tag{D3}$$

then the two poles z_+ and z_- will in our applications usually occur in the upper and lower half-planes respectively so that

$$J = F/C + (i/4)(B^2-AC)^{-1/2}\{(D+2E/z_+ + F/z_+^2)\times \\ \times[1-\exp(2iz_+)] + (D+2E/z_- + F/z_-^2)\times \\ \times[1-\exp(-2iz)]\} \tag{D4}$$

where

$$z_+ = [-B+(B^2-AC)^{1/2}]/A \tag{D5}$$

$$z_- = [-B-(B^2-AC)^{1/2}]/A \tag{D6}$$

For the special case of *no drift*, $B = E = 0$, we get

$$J = \frac{F}{C} - \frac{1}{2}\left(\frac{A}{C}\right)^{1/2}\times \\ \times\left(\frac{F}{C}-\frac{D}{A}\right)\{1-\exp[-2(C/A)]^{1/2}\} \tag{D7}$$

In our applications, we always find $|\exp(-2iz_-)| \ll 1$. When drift is not predominant, or at sufficiently high frequencies in any case, we have $|\exp(2iz_+)| \ll 1$, i.e. when *boundary interference effects* are negligible:

$$J \approx (F/C + \tfrac{1}{2}(AC-B^2)^{1/2}\times \\ \times[D-2(EB/C)-F(AC-2B^2)/C^2] \tag{D8}$$

When drift is strong, we must retain interference effects, but can *neglect diffusion*, i.e. with $D = A = 0$

$$J = (F/C) \pm i[(FB-EC)/C^2][1-\exp(\mp iC/B)] \tag{D9}$$

where the sign is always chosen so that the magnitude of the exponential is less than unity. For extreme drift $E \to \infty$, $B \to \infty$ but E/B finite

$$J \approx E/B \tag{D10}$$

APPENDIX E

Exact solution, trapping plus diffusion

Equation (5.16) can be rewritten in matrix form

$$\frac{\partial}{\partial t}\begin{pmatrix}\Delta n(k,t)\\ \Delta\hat{n}(k,t)\end{pmatrix} = -\mathbf{\Lambda}(k)\begin{pmatrix}\Delta n\\ \Delta\hat{n}\end{pmatrix} \tag{E1}$$

where

$$\mathbf{\Lambda}(k) = \begin{pmatrix}\nu+r+iv_p k+Dk^2, & -g'+i\hat{v}_p k+\hat{D}_p k^2\\ -\nu, & g'\end{pmatrix} \tag{E2}$$

is the decay matrix. Equation (5.14) has the matrix Green's function

$$\mathbf{K}(x,x',t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp[ik(x-x')]\exp[-\mathbf{\Lambda}(k)t]\, dk \tag{E3}$$

The autocorrelation function and its Laplace transform are now matrices:

$$\mathbf{\Phi}(t) = \frac{1}{L}\int_0^L\int_0^L dx\, dx' K$$

$$= \frac{L}{2\pi}\int_{-\infty}^{\infty}\left[\frac{\sin(kL/2)}{(kL/2)}\right]^2\exp[-\mathbf{\Lambda}(k)t]\, dk \tag{E4}$$

$$\mathbf{g}(\omega) = Re\frac{1}{\pi}\int_{-\infty}^{\infty}\left(\frac{\sin z}{z}\right)^2\frac{1}{i\omega+\mathbf{\Lambda}(z)}\, dz \tag{E5}$$

The effect of diffusion is therefore to replace the matrix $[i\omega+\mathbf{\Lambda}(0)]^{-1}$ by $\mathbf{g}(\omega)$. We may make this replacement in the first form of equation (2.21) to obtain the exact result for the noise spectrum. (The approximations used in the text are essentially equivalent to taking account of the change produced by diffusion only in $g_{22}(\omega)$. The other terms are either small compared to this dominant term or not significantly affected by diffusion.) For the exact

VOL 14 196

result it is convenient to add the terms in equation (2.21) underneath the integral sign. The result is a single integral of the form evaluated in Appendix D.

REFERENCES

1. McWhorter A. L., *Semiconductor Surface Physics* (edited by Kingston R. H.) pp. 207–228. University of Pennsylvania Press, Philadelphia (1957).
2. Brophy J. J., *J. Appl. Phys.* **29,** 1377L (1958); Maple T. G., Bess L. and Gebbie H. L., *J. Appl. Phys.* **26,** 490 (1955); Montgomery H. C., *Bell Syst. Tech. J.* **31,** 950 (1952).
3. Montgomery H. C., *Bell Syst. Tech. J.* **31,** 950 (1952); Brophy J. J. and Rostoker M., *Phys. Rev.* **100,** 754 (1955); Brophy J. J., *Phys. Rev.* **106,** 675 (1957); Bess L., *J. Appl. Phys.* **26,** 1377 (1955).
4. MacRae A., Thesis, Physics Dept., Syracuse University (1959).
5. Rollin B. V. and Templeton I. M., *Proc. Phys. Soc. Lond.* B**66,** 259 (1953); Firle T. E. and Winston H., *Bull. Amer. Phys. Soc.* **30,** 44 (1955).
6. McWhorter A. L., Ref. 1; van der Ziel A., *Physica, s'Grav.* **16,** 359 (1950); du Pre F. K., *Phys. Rev.* **78,** 615 (1950).
7. Hornbeck J. A. and Haynes J. R., *Phys. Rev.* **97,** 311 (1955); **100,** 606 (1955).
8. van Roosbroeck W., *Bell Syst. Tech. J.* **39,** 515 (1960); *Phys. Rev.* (to be published).
9. Lax M., *Rev. Mod. Phys.* **32,** 25 (1960), see especially Section 11. Also Hill J. E. and van Vliet K. M., *Physica, s'Grav.* **24,** 709 (1958).
10. Chandrasekhar S., *Rev. Mod. Phys.* **15,** 1 (1943), see Chapter 3.
11. Watson G. N., *Theory of Bessel Functions* (2nd. Ed.), Watson's lemma, Section 8.3, p. 236. Cambridge University Press (1948).
12. Champlin K. S., *Study of Noise in Semiconductors, Quart. Rep. Elect. Engng. Dep. Univ. Minnesota,* March 1 (1959) and June 1 (1959). (unpublished).
13. Geballe T. H., *Semiconductor Chemistry* (edited by Hannay N. B.) Chapter 8, equations (8.24), (8.27). Reinhold, New York (1959).
14. Haynes J. R., Lax M. and Flood W., *J. Phys. Chem. Solids* **8,** 392 (1959), Fig. 6.
15. Morse P. M. and Feshbach H., *Methods of Theoretical Physics,* p. 832, equation (7.2.68). McGraw-Hill, New York (1953).
16. Morse P. M. and Feshbach H., Ref. 15, p. 821, equation (7.2.39).
17. Landau L. and Lifshitz E. M., *Statistical Physics,* Section 74. Addison-Wesley, Reading, Mass. (1958); Debye P. and Hückel E., *Phys. Z.* **24,** 185 (1923).
18. Mayer J. E., *J. Chem. Phys.* **18,** 1426 (1950) uses a classical summation of ring diagrams.

VOL. 14 1960

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 268–270. Printed in Great Britain.

SURFACE ELASTIC WAVES IN SEMICONDUCTORS

D. C. GAZIS and R. HERMAN

Research Laboratories, General Motors Corporation, Warren, Michigan

and

R. F. WALLIS

U.S. Naval Research Laboratory, Washington 25, D.C.

1. INTRODUCTION

AN understanding of the nature of surface elastic waves in semiconductors is important in the study of such phenomena as electrical conductivity and the scattering and absorption of radiation at surfaces and interfaces. The existence of surface waves in isotropic elastic media has been known since the pioneering work of LORD RAYLEIGH[1]. These surface waves, known as Rayleigh waves, are characterized by displacement amplitudes which decrease exponentially with increasing distance from the free surface.

There have been relatively few treatments of surface waves in anisotropic media such as crystals. STONELEY[2] has studied the case of cubic crystals and has found that Rayleigh-type surface waves exist for some but not for all physically allowable values of the three elastic constants. STONELEY assumed that the displacement components (u, v, w) at a point in the crystal can be written in the form

$$(u, v, w) = (U, V, W) \exp \kappa[-qz+i(lx+my-ct)] \tag{1}$$

where c is the phase velocity, κ the wave number, q the attenuation constant and U, V, W are constants. The free surface is assumed to be defined by $z = 0$ and l, m are the direction cosines for the direction of wave propagation parallel to the surface.

It is necessary that the displacements satisfy both the equations of motion and the boundary conditions at the free surface. For given values of c, l and m there are at most three values of q which have positive real parts and for which equation (1) satisfies the equations of motion. By substituting these three values of q into equation (1) and taking a linear combination of the three displacements thus obtained, one can satisfy the boundary conditions provided the phase velocity c is properly chosen. If the three values of q turn out to be positive and real, then a surface wave exists and it is a Rayleigh-type wave. It sometimes happens, however, that two of the q's are complex conjugates and the third is positive real. A surface wave again exists, but it is not of the Rayleigh-type and we shall designate it by the term generalized Rayleigh wave. Finally, if one or more of the q's is pure imaginary, no surface wave exists.

2. SURFACE WAVES IN DIAMOND, SILICON, GERMANIUM AND LEAD SULFIDE

In a recent investigation, GAZIS *et al.*[3] have determined the ranges of the elastic constants for which the various types of surface waves exist on the (001) surface of a cubic crystal. The results are shown in Fig. 1 where the elastic constant ratios $g_1 = c_{11}/c_{44}$ and $g_2 = c_{12}/c_{44}$ are plotted on the x- and y-axes, respectively. The region above the line $g_1+2g_2 = 0$ and below the line $g_1-g_2 = 0$ corresponds to a stable crystal and is physically acceptable. If the elastic constants of a crystal fall in the region below the curve marked bound (1) and above the line $g_1+2g_2 = 0$, then Rayleigh waves exist at (001) surfaces. The position of bound (1) depends to a slight extent on the direction of propagation of the surface wave, the curve shown being that for the (100) direction.

If the elastic constants fall below the curve marked bound (2) but above bound (1), surface waves again exist on a (001) surface, but they are now of the generalized Rayleigh-type. This is true for all directions of propagation from (100) to (110).

VOL. 14 1960

Above bound (2) but below the line $g_1-g_2=0$, there is a range of directions of propagation about the (110) direction for which no surface waves can exist in a (001) surface. Under these circumstances at least one of the attenuation constants q turns out to be pure imaginary so that the wave does not damp out away from the surface.

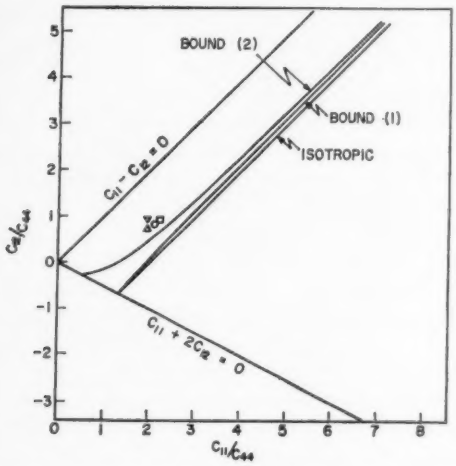


FIG. 1. Bounds for Rayleigh waves (bound 1) and for generalized Rayleigh waves (bound 2) in terms of the elastic constant ratios c_{11}/c_{44} and c_{12}/c_{44}. □: diamond, ○: silicon, △: germanium, ▽: lead sulfide.

Using values of room temperature elastic constants taken from the *American Institute of Physics Handbook*, pp. 2–56, one finds that the points for diamond, silicon, germanium and lead sulfide lie above bound (2) in Fig. 1. These materials, therefore, possess generalized Rayleigh waves in some but not all directions of propagation on a (001) surface. We have carried out calculations of the phase velocities and attenuation constants for surface waves propagating in various directions on the (001) surface of the above mentioned semiconductors. The numerical work was carried out on the IBM-704 digital computer at the Research Laboratories of the General Motors Corporation.

The results for the attenuation constants for silicon are presented in Fig. 2 where the real attenuation constant q_1, and the real and imaginary parts q_r and q_i of the two complex conjugate attenuation constants are plotted against the angle ϕ between the direction of propagation of the surface wave and the (100) axis. The real attenuation constant q_1, has a maximum at $\phi=0$ and decreases monotonically to zero at $\phi\simeq 43°$. For ϕ in the range $\sim 43°<\phi<45°$ no surface wave exists. For $\phi=45°$, however, a surface wave again exists because the solution corresponding to q_1 is no longer needed to satisfy the boundary conditions.

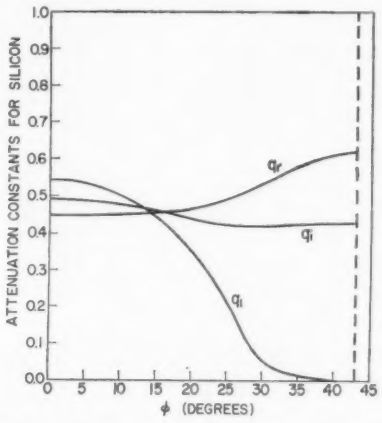


FIG. 2. Attenuation constants for silicon as a function of the angle between the direction of propagation and the (100) axis.

The values of q_r and q_i are relatively insensitive to changes in ϕ.

In ordinary physical units the attenuation constants are equal to the products κq where κ is inversely proportional to the wavelength of the surface wave. Consequently, the shorter the wavelength the more rapidly the wave attenuates away from the surface.

The attenuation constants have also been calculated for diamond, germanium and lead sulfide. In each case the results are very similar to those for silicon given in Fig. 2. In particular, the real attenuation constant q_1 becomes zero for each material when ϕ is very nearly equal to 43°.

The results for the phase velocities are presented in terms of the ratio p of the surface wave phase velocity c to $(c_{44}/\rho)^{1/2}$ where ρ is the density. The quantity $(c_{44}/\rho)^{1/2}$ is the phase velocity for transverse bulk waves propagating in the (100) direction. A plot of p as a function of the angle between the direction of propagation of the surface wave and the (100) axis is given in Fig. 3. For all four materials considered the surface wave phase

velocity increases slowly to a maximum and then decreases as the direction of propagation varies from (100) toward (110). In general the value of p is less than unity so that the surface waves have a lower phase velocity than do the transverse bulk waves. The total variation of p with ϕ for any one of the four semiconductors is quite small and does not exceed about 10 per cent.

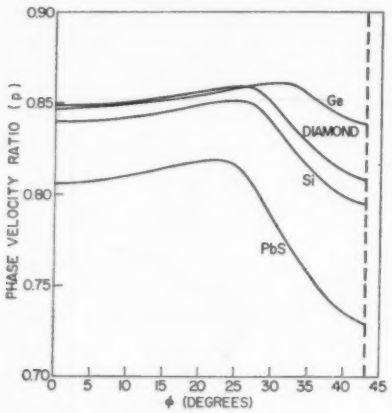


FIG. 3. Phase velocity ratio as a function of the angle between the direction of propagation and the (100) axis.

3. COMMENTS ON THE LATTICE THEORY OF SURFACE WAVES

Several investigations have been made[3–5] on the theory of surface waves in crystals from the lattice vibrational point of view. It has been found[3, 4] that under proper circumstances surface waves are associated with the accoustical branch which become identical with Rayleigh waves or generalized Rayleigh waves in the limit of long wavelengths. As the wavelength becomes shorter and comparable to the lattice spacing, these acoustic surface waves exhibit dispersion,[3] that is, the phase velocity decreases with decreasing wavelength.

It has also been found[4, 5] that surface lattice vibrational waves may be associated with the optical branch. These optical surface waves may have frequencies which lie in the "forbidden" gap between the usual acoustical and optical branches.

Carrying out a lattice vibrational theory of surface waves in the semiconductors considered here is somewhat difficult for several reasons. First, these materials possess more than one atom per unit cell and second, a completely satisfactory model may involve a great many interatomic interactions. The calculations, expecially the satisfying of the proper boundary conditions, may therefore become very complicated. The problem is by no means hopeless, however, and it is believed that lattice vibrational calculations can lead to worthwhile results.

REFERENCES

1. RAYLEIGH LORD, *Proc. Lond. Math. Soc.* **17,** 4 (1885).
2. STONELEY R., *Proc. Roy. Soc.* A**232,** 447 (1955).
3. GAZIS D. C., HERMAN R. and WALLIS R. F., to be published.
4. LIFSHITZ I. M. and ROSENZWEIG L. N., *Zh. eksp. teor. fiz.* **18,** 1012 (1948).
5. WALLIS R. F., *Phys. Rev.* **105,** 540 (1957).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 271–290. Printed in Great Britain.

GALVANOMAGNETIC PROPERTIES OF SEMICONDUCTOR THIN FILMS AND SURFACE LAYERS

A. AMITH

RCA Laboratories, Princeton, New Jersey

VOL. 14 1960

INTRODUCTION

THE purpose of this paper is to study the behaviour of carriers in thin films as well as in surface layers* of semiconductors, under the influence of electric and magnetic fields. The reason for the treatment of both thin films (wherein no space charge is assumed to exist) and surface layers lies in the marked resemblance of these two physical situations. Those charge carriers which are attracted to the surface by an electrostatic potential (which is due to the surface electric field) are constrained to move along the surface in a potential well. This well is, in effect, a thin film bounded by the actual surface on one side and by the potential incline on the other. Unlike the carriers in the case of the thin films, however, the carriers in the surface potential well possess electrostatic energy which is position dependent; furthermore, the greater the total energy of those carriers, the thicker the effective film in which they are constrained to move (this is due to the dependence of the potential upon distance from the surface), and carriers of sufficiently high energy are not bound in the well at all, and can move anywhere in the entire body of the sample. A film, or surface layer, is considered to be "*thin*" when its thickness is comparable to (or smaller than) the mean free path of the carriers.

The emergence of a "size effect", which is the variation of the galvanomagnetic properties with the thickness of the film (exclusive of the effects of the charge concentration or depletion by the surface potential), comes about as a result of the additional scattering mechanism which now affects an increasing fraction of the carriers: *surface scattering*. In the geometry considered here, the surface (or two surfaces) is parallel to the current flow and normal to the applied magnetic field (which is along the layer's thin dimension). There are two distinct (though not mutually exclusive) manners in which carriers can be scattered from the surface. In the case of *specular* reflection only the direction of the carrier's momentum normal to the current flow is reversed. Clearly none of the transport properties would be altered, and the properties of the thin layer in this case would be indistinguishable from those of the bulk (except for the effects of an electrostatic potential, which are not considered to be "size effects"). When the carriers undergo *diffuse* scattering at the surface, Knudsen's cosine law obtains and the density of scattered carriers is isotropic. This corresponds to termination of all free paths and loss of the carriers' velocities after they have been scattered. It is the diffuse scattering which is our main concern.

The decreased conductivity of thin metallic films was discussed in detail by FUCHS[1], and his treatment was extended to include the effects of a transverse magnetic field by SONDHEIMER[2]. SONDHEIMER predicted that the conductivity should undergo strongly damped oscillations as a function of the magnetic field; these were observed for the first time in a thin sodium wire at 1°K by BABISKIN and SIEBENMANN[3]. In the case of a surface layer wherein a space charge exists because of charges in surface states or an external electric field, the galvanomagnetic properties differ from those in the bulk due to two causes: the scattering which the carriers undergo at the surface, and the change from the bulk value of the carriers' densities due to the surface electrostatic field. The effect of scattering from the surface upon the conduction in a space-charge layer of a semiconductor was first considered by SCHRIEFFER[4] whose method

* The surface layer is that region of the bulk material, adjacent to the surface, in which a net space charge exists.

was extended to include a magnetic field by ZEMEL[5]. ZEMEL and PETRITZ[6] performed experiments on the magnetoconductivity and Hall effect of surface layers in near intrinsic germanium, following the development of a formalism by the second author.[7]

In the present study, the Boltzmann transport equation is solved for the case when a space charge is present, and the thin film (without a space charge) represents but a special limiting form. It is noteworthy that the steady state treatment which follows is applicable even when the electrostatic potential in the surface layer is altered in a "field effect" experiment[8] wherein changes in local carrier densities take place; this is due to the large difference between the sundry scattering relaxation times and the lifetimes of the free excess carriers. The problem can be separated conveniently in terms of independent distribution functions, identical in form, for holes and for electrons. The spatial dependence of the potential is known as a function of the electrostatic field at the surface from the solution of Poisson's equation, and it matters nought whether the field at the surface is due to charges in surface states, to an external field (applied capacitively) or to both. The specific model adopted in the present discussion is based upon the assumption of spherical energy surfaces and the existence of a relaxation time which is independent of energy, isotropic, and the same for all carriers of one kind. The effective mass is taken to be a scalar quantity. Furthermore, the semiconductor is non-degenerate in the bulk as well as in the surface space-charge region. A linear surface potential which is truncated when it reaches the bulk potential was used for the computations.†

The measurable physical quantities of interest are the average (over the thickness) longitudinal current density and the Hall field. From these one can derive the conductivity, the Hall constant and the magnetoresistance. Further, it will be of interest to define effective mobilities which would differ from the bulk mobilities due to the diffuse scattering from the surface. These effective mobilities will depend upon both the thickness of the layer and the value of the surface field. The effective mobilities can be defined in a number of alternate fashions: from the electric current in the thin layer,[4] from the Hall coefficient or from the magnetoresistance.[7] The choice depends upon the physical description which one prefers. In the present work three types of conductivity effective mobilities are defined: the mobility derived from the total current flowing in the thin layer, the mobility of the excess (due to the surface potential) carriers alone, and the mobility of the excess carriers from which has been subtracted the surface scattering which is independent of the surface potential. The latter achieves (for the case of low magnetic fields) a formal separation of the scattering in the presence of a surface potential, from that which exists in its absence. Furthermore, in addition to the field effect mobility, there are defined two types of effective Hall mobilities: one involves averaging over the mobility itself, and the other over its square.

THE DISTRIBUTION FUNCTIONS

The Maxwell–Boltzmann thermal equilibrium distribution function has the following familiar form for a single type of carrier

$$f_0 = n_0(m^*/2\pi kT)^{3/2} \times \\ \times \exp\left[-(\tfrac{1}{2}m^*v^2+q\psi)/kT\right], \qquad (1)$$

where q is the carrier's electric charge, m^* is its effective mass, n_0 is the thermal equilibrium density of carriers in the bulk‡ and ψ is the electrostatic potential which may exist near the surface ($\psi = 0$ in the bulk). The local densities of holes and of electrons are given by

$$p = p_0 \exp(-e\psi/kT), \qquad n = n_0 \exp(e\psi/kT). \quad (2)$$

It is wont to describe the solution of the Boltzmann transport equation[9] as

$$f = f_0 + f_1. \qquad (3)$$

In the customary ohmic approximation (applied to the drift (E_x, 0, 0) and the Hall (0, E_y, 0) field), the transport equation which has to be satisfied

† Such a model does not predict any bulk magnetoresistance in the case of a one-carrier semiconductor. However, the diffuse scattering at the surface reduces the mobilities of some of the carriers and consequently gives rise to magnetoresistance.

‡ When dealing with a single type of carrier, n will refer to its local density. When dealing with electrons and holes, n will refer to the electron density.

by f_1 in the presence of both electric and magnetic fields, becomes:

$$\frac{f_1}{\tau v_z}+\frac{\partial}{\partial z}f_1+\frac{qE_z}{m^*v_z}\frac{\partial}{\partial v_z}f_1+$$
$$+\frac{qH_z}{m^*v_z c}\left[v_y\frac{\partial}{\partial v_x}f_1-v_x\frac{\partial}{\partial v_y}f_1\right]$$
$$=\frac{qf_0}{kTv_z}[E_xv_x+E_yv_y]. \tag{4}$$

The general solution of this quasi-linear partial differential equation is obtained by the method of LAGRANGE[10]:

$$f_1=e^{-\mathscr{L}}\left\{D+\frac{q\tau v f_0}{kT(1+\beta^2)}e^{\mathscr{L}}[E_x(\cos\xi-\beta\sin\xi)+\right.$$
$$\left.+E_y(\sin\xi+\beta\cos\xi)]\right\}, \tag{5}$$

in which

$$v_x=v\cos\xi, \qquad v_y=v\sin\xi, \tag{6a}$$

$$\mathscr{L}(\tilde{\epsilon},z)=\int_{z_p}^{z}\frac{dz}{\tau[2(\tilde{\epsilon}-q\psi)/m^*]^{1/2}}, \tag{6b}$$

$$\tilde{\epsilon}=\tfrac{1}{2}m^*v_z^2+q\psi(z)\equiv q\psi(z_p), \tag{6c}$$

and

$$\beta\equiv\frac{q\tau}{m^*c}H_z. \tag{6d}$$

D is a constant in the variable $\mathscr{L}$, and its value will be determined by introducing the appropriate boundary conditions at the surface.

Diffuse Scattering

The distribution function for those carriers which are leaving the surface having been diffusely scattered thereat, should be independent of direction altogether: f_1 must vanish there. At the $z=0$ surface the perturbation function f_1 has to vanish for carriers with $v_z>0$, while at the $z=d$ surface the perturbation function f_1 has to vanish for carriers with $v_z<0$. Clearly, two analytic forms of the distribution function are needed in order to describe the system. $f_1^>$ will describe the carriers in that half of momentum space where $v_z>0$, and $f_1^<$ will describe the carriers in the other half where $v_z<0$.

If there exists a surface potential near the $z=0$ surface (and the direction of the space charge field is the same throughout the sample), the situation depends upon whether the potential is attractive or repulsive. (1) If the potential is *attractive*, then there will be carriers constrained to move in the potential well bounded by $z=0$ and by the potential incline at $z=z_p$; for those carriers $f_1^>$ will be determined by the condition that it vanish at $z=0$, and $f_1^<$ will be determined by the condition of specular reflection at the potential incline. For those carriers with energies high enough so that they are not bound, $f_1^>$ will vanish at $z=0$ and $f_1^<$ will vanish at $z=d$. (2) If the potential is *repulsive*, then again one needs to distinguish between the bound carriers which are prevented by the barrier from reaching the $z=0$ surface, and the higher energy carriers which can reach it. For the bound carriers, $f_1^>$ is determined by specular reflection at the potential incline, and $f_1^<$ vanishes at $z=d$. For the unbound carriers, as in the case of the attractive potential, $f_1^>=0$ at $z=0$ and $f_1^<=0$ at $z=d$. (3) When we are dealing with a sample wherein $\psi=0$ throughout, that is, a *thin film* without any space charge, then simply $f_1^>=0$ at $z=0$ and $f_1^<=0$ at $z=d$.

In studying the properties of the surface layer adjacent to $z=0$, we shall usually assume that the sample is thick in comparison with the carriers' mean free paths, so that $f_1^<$ in the region of interest will be simply the bulk function $f_1^{0\beta}$ (*vide infra*) for all the unbound carriers as well as for the bound carriers in the case of the repulsive potential.

Attractive potential. $f_1^>(v_z)$ is determined by choosing the constant D in equation (5) so that f_1 is zero at $z=0$.

Define

$$\mathscr{L}_0\equiv\mathscr{L}(\tilde{\epsilon},z=0) \text{ and } \Delta\mathscr{L}\equiv\mathscr{L}_0-\mathscr{L}$$

consequently†

$$f_1^>(v_z)=f_1^{0\beta}-\frac{q\tau v f_0}{kT(1+\beta^2)}\times$$
$$\times\exp(\Delta\mathscr{L}^>)\{X\cos(\beta\Delta\mathscr{L}^>)+\Upsilon\sin(\beta\Delta\mathscr{L}^>)\}, \tag{7}$$

† Since the sense of v_z has not been preserved explicitly in the new variable $\mathscr{L}$, the latter was designated in equation (7) as $\mathscr{L}^>$, to indicate that it refers to subsequent integration over the positive v_z subspace. The positive sign of the square root in the integrand will apply to $\mathscr{L}^>$. $\mathscr{L}^<$ will refer to the negative v_z subspace, and the negative sign of the square root will apply to it.

VOL. 14 1960

wherein

$$X = (E_x+\beta E_y)\cos\xi+(E_y-\beta E_x)\sin\xi,$$
$$\Upsilon = (E_x+\beta E_y)\sin\xi-(E_y-\beta E_x)\cos\xi, \quad (8)$$

and $f_1^{0\beta}$ is the bulk form of the function f_1:

$$f_1^{0\beta} = \frac{q\tau f_0}{kT(1+\beta^2)}\{(E_x+\beta E_y)v_x+(E_y-\beta E_x)v_y\}. \quad (9)$$

Carriers moving away from the surface toward the potential incline are described by $f_0+f_1^{>}(v_z)$, and those moving toward the surface are described by $f_0+f_1^{<}(v_z)$. The latter are, however, the very carriers which were reflected specularly at the potential incline, that is: $f_0+f_1^{<}(v_z) = f_0+f_1^{>}(-v_z)$. Hence, at the potential incline:

$$f_1^{<}(v_z) = f_1^{>}(-v_z). \quad (10)$$

For the unbound carriers, $f_1^{>}(v_z)$ is identical with the function of equation (7), and $f_1^{<}(v_z)$ is identical in form with the function of equation (7) in which $\Delta\mathscr{L}^{>}$ has been replaced by $\Delta\mathscr{L}^{<}$, where $\Delta\mathscr{L}^{<} \equiv \mathscr{L}^{<}(\bar{\epsilon}, z = d)-\mathscr{L}^{<}$. Usually this function $f_1^{<}(v_z)$ in the space-charge region will be simply $f_1^{0\beta}$ unless the sample's thickness d is of the order of a mean free path.

Repulsive potential. $f_1^{>}(v_z)$ and $f_1^{<}(v_z)$ for the unbound carriers are identical in form with the same functions for the unbound carriers in the case of an attractive potential. For the bound carriers which cannot reach the $z = 0$ surface, $f_1^{>}(v_z)$ is determined by specular reflections at the potential incline of carriers arriving there from the $z = d$ surface. For thick samples, however, $f_1^{>}(v_z)$ and $f_1^{<}(v_z)$ for the bound carriers will be simply the bulk function $f_1^{0\beta}$.

No surface potential. The boundary conditions which determine the explicit forms of equation (5) in this case are that $f_1^{>}$ vanish at the $z = 0$ surface, and that $f_1^{<}$ vanish at the $z = d$ surface. $f_1^{>}$ is obtained from equation (7) after setting $\psi = 0$, and $f_1^{<}$ is obtained from a similar equation applied at the other surface. Alternatively, one could have proceeded directly from equation (4) in which the term containing E_z had been eliminated. The two functions are:

$$f_1^{>}(v_z) = f_1^{0\beta} - \frac{q\tau v f_0}{kT(1+\beta^2)}\exp(-z/\tau v_z)\left\{X\cos\frac{\beta z}{\tau v_z}-\Upsilon\sin\frac{\beta z}{\tau v_z}\right\}, \quad (12a)$$

$$f_1^{<}(v_z) = f_1^{0\beta} - \frac{q\tau v f_0}{kT(1+\beta^2)}\exp(d-z/\tau v_z)\left\{X\cos\beta\frac{d-z}{\tau v_z}+\Upsilon\sin\beta\frac{d-z}{\tau v_z}\right\}. \quad (12b)$$

Combination of equations (7) and (10) yields for the bound carriers*:

$$f_1^{<}(-v_z) = f_1^{0\beta}(1+e^{\mathscr{L}}) - \frac{q\tau v f_0}{kT(1+\beta^2)}e^{\mathscr{L}}\{X[e^{\mathscr{L}^{>}_0}\cos(\beta\mathscr{L}^{>}_0)+\cos(\beta\mathscr{L})]+$$
$$+\Upsilon[e^{\mathscr{L}^{>}_0}\sin(\beta\mathscr{L}^{>}_0)+\sin(\beta\mathscr{L})]\}. \quad (11)$$

CURRENTS AND COEFFICIENTS

The new distribution functions give rise to electrical current densities in both the x direction (drift current) and the y direction (Hall current):

$$J_x = q\iiint v_x f_1\, dv_x\, dv_y\, dv_z,$$
$$J_y = q\iiint v_y f_1\, dv_x\, dv_y\, dv_z. \quad (13)$$

f_1, and consequently the current densities, will depend in general upon z — the distance from the surface. In order to develop expressions for quantities which can be compared with experiment, average current densities will be defined for

* For the cases in which a surface potential is present, we write down $f_1^{<}(-v_z)$ rather than $f_1^{<}(v_z)$, because $f_1^{<}(-v_z)$ can be subsequently integrated over the same positive v_z subspace as $f_1^{>}(v_z)$.

a layer of thickness d adjacent to the $z = 0$ surface:

$$\overline{J_x} = d^{-1}\int_0^d J_x\,dz, \qquad J_y = d^{-1}\int_0^d J_y\,dz, \tag{14}$$

where

$$\int_0^d J_x\,dz$$

and

$$\int_0^d J_y\,dz$$

are the currents per unit width of the surface area, flowing in the thickness d. The experimental condition is that there be no net current in the y direction. This yields the relationship between the Hall field and the applied magnetic and electric drift fields. Subsequently explicit expressions are obtained for the longitudinal current as a function of the applied fields, for the Hall constant and for the magnetoresistance.

Diffuse scattering

The distribution functions for the present case are not given in terms of the variables v_x, v_y, v_z and z but rather in terms of the variables v, ξ, and $\mathscr{L}$. Rather than transform the distribution functions to the old variables, it would be simpler to transform the integration itself to the new variables. Although $\mathscr{L}$ and $\tilde{\epsilon}$ do not constitute an orthogonal subsystem, the transformation of the integration space is straightforward. The Jacobian of the transformation is

$$\frac{\partial(v_x\ v_y\ v_z\ z)}{\partial(v\ \xi\ \mathscr{L}\ \tilde{\epsilon})} = \frac{\tau v}{m^*}. \tag{15}$$

The average current densities in terms of the distribution functions of equations (7) and (11) are:

$$\overline{J_x} = \frac{q}{m^* d}\int d\tilde{\epsilon}\int d\mathscr{L}\int_0^{2\pi} d\xi\int_0^{\infty} dv\,\tau v^2 f_1\cos\xi, \tag{16}$$

$$\overline{J_y} = \frac{q}{m^* d}\int d\tilde{\epsilon}\int d\mathscr{L}\int_0^{2\pi} d\xi\int_0^{\infty} dv\,\tau v^2 f_1\sin\xi. \tag{16}$$

While the integrations over v and ξ are obvious, the integrations over $\tilde{\epsilon}$ and $\mathscr{L}$ have to be done separately for the bound and the unbound carriers in both cases of the attractive and repulsive potentials.

It is possible to write the average current densities in the following symmetrical form:

$$\overline{J_x} = E_x\frac{\bar{\sigma}}{1+\beta^2}(1-A)+\beta E_y\frac{\bar{\sigma}}{1+\beta^2}(1-B), \tag{17}$$

$$\overline{J_y} = E_y\frac{\bar{\sigma}}{1+\beta^2}(1-A)-\beta E_x\frac{\bar{\sigma}}{1+\beta^2}(1-B),$$

where

$$\bar{\sigma} = \bar{n}e\mu, \qquad \bar{n} = n_0 S/d,$$

and

$$S = \int_0^d dz\epsilon xp\{-q\psi/kT\}, \tag{18}$$

where d is the extent of the surface space charge layer.

The various galvanomagnetic quantities assume the following forms for the single carrier case:

$$E_y = \beta E_x\frac{1-B}{1-A},$$

$$\overline{J_x} = \frac{\bar{\sigma}}{1+\beta^2}E_x\left\{1-A+\beta^2\frac{(1-B)^2}{1-A}\right\},$$

$$R = \frac{1}{\bar{n}qc}\frac{(1+\beta^2)(1-B)}{(1-A)^2+\beta^2(1-B)^2}, \tag{19}$$

$$M = 1-\frac{1-A+\beta^2(1-B)^2/(1-A)}{(1+\beta^2)(1-A^0)}.$$

For both electrons and holes:

$$E_y = \frac{H}{c}E_x\frac{\Omega}{\Pi},$$

$$\overline{J_x} = \frac{eE_x}{(1+\beta_h^2)(1+\beta_e^2)\Pi}\left\{\Pi^2+\left(\frac{H}{c}\Omega\right)^2\right\}, \tag{20a}$$

$$R = \frac{(1+\beta_h^2)(1+\beta_e^2)}{ec}\frac{\Omega}{\Pi^2+\left(\frac{H}{c}\Omega\right)^2},$$

where

$$\Pi = \bar{p}\mu_h(1+\beta_e^2)(1-A_h)+\bar{n}\mu_e(1+\beta_h^2)(1-A_e),$$

$$\Omega = \bar{p}\mu_h^2(1+\beta_e^2)(1-B_h)-\bar{n}\mu_e^2(1+\beta_h^2)(1-B_e). \tag{20b}$$

In the limit of small magnetic fields, the expressions are:

$$E_y = \frac{H}{c} E_x \frac{\tilde{p}\mu_h^2(1-B_h^0)-\tilde{n}\mu_e^2(1-B_e^0)}{\tilde{p}\mu_h(1-A_h^0)+\tilde{n}\mu_e(1-A_e^0)},$$

$$\overline{J_x} = eE_x\{\tilde{p}\mu_h(1-A_h^0)+\tilde{n}\mu_e(1-A_e^0)\}, \quad (21)$$

$$R = \frac{1}{ec}\frac{\tilde{p}\mu_h^2(1-B_h^0)-\tilde{n}\mu_e^2(1-B_e^0)}{[\tilde{p}\mu_h(1-A_h^0)+\tilde{n}\mu_e(1-A_e^0)]^2}.$$

In the present situation there is great difference between the roles played by the two kinds of carriers participating in the conduction in the channel: one of them is in an attractive surface potential while the other is on a repulsive barrier. It will be shown subsequently that the scattering correction factors A and B are quite different for the two cases. In practice it will be necessary to consider the two carrier model only for some cases of inversion layers.

Attractive potential. In this case $q\psi < 0$, and it is instructive to write the scattering correction factors as follows:

$$A_a = A_{a(\text{bound})} + A_{a(\text{unbound})},$$

$$B_a = B_{a(\text{bound})} + B_{a(\text{unbound})}. \quad (22)$$

The correction factors $A_{a(\text{bound})}$ and $B_{a(\text{bound})}$ pertain to the bound carriers which are constrained to move in the attractive potential well and which are scattered diffusely at the $z = 0$ surface and reflected specularly at the potential incline. The correction factors $A_{a(\text{unbound})}$ and $B_{a(\text{unbound})}$ pertain to the unbound carriers which are scattered diffusely at the $z = 0$ surface but can move anywhere in the thick samples.

Heretofore[4,5], the contribution of the unbound carriers has always been neglected.† Their contribution is significant, as could be surmised, only at low values of the surface potential. The ratios $A_{a(\text{unbound})}/A_a$ and $B_{a(\text{unbound})}/B_a$ fall off exponentially as a function of $q\psi/kT$.

The scattering correction factors assume the following specific forms:

$$A_{a\,(\text{bound})} = \frac{1}{\kappa_a 2\pi^{1/2}(1+\beta^2)}\Big\{(3\beta^2-1)I_{1,\mathscr{R}}+(\beta^3-3\beta)I_{1,\mathscr{I}}-$$
$$-(1+\beta^2)I_{2,\mathscr{R}}-(\beta+\beta^3)I_{2,\mathscr{I}}+(1+\beta^2)I_1^{\beta=0}+(1-3\beta^2)I_0^{\beta=0}\Big\}, \quad (23a)$$

$$A_{a\,(\text{unbound})} = \frac{1}{\kappa_a 2\pi^{1/2}(1+\beta^2)}\int_{\epsilon_a}^{\infty} d\epsilon e^{-\epsilon}[(\beta^2-1)\exp(\Delta\mathscr{L}_{\vec{d}}^{>})\cos(\beta\Delta\mathscr{L}_{\vec{d}}^{>})-$$
$$-2\beta\exp(\Delta\mathscr{L}_{\vec{d}}^{>})\sin(\beta\Delta\mathscr{L}_{\vec{d}}^{>})+1-\beta^2], \quad (23b)$$

$$B_{a\,(\text{bound})} = \frac{1}{\kappa_a 2\pi^{1/2}(1+\beta^2)}\Big\{(\beta^2-3)I_{1,\mathscr{R}}+\frac{1-3\beta^2}{\beta}I_{1,\mathscr{I}}-(1+\beta^2)I_{2,\mathscr{R}}+\frac{1+\beta^2}{\beta}I_{2,\mathscr{I}}+$$
$$+(1+\beta^2)I_1^{\beta=0}+(3-\beta^2)I_0^{\beta=0}\Big\}, \quad (24a)$$

$$B_{a\,(\text{unbound})} = \frac{1}{\kappa_a 2\pi^{1/2}(1+\beta^2)}\int_{\epsilon_a}^{\infty} d\epsilon e^{-\epsilon}[2-2\exp(\Delta\mathscr{L}_{\vec{d}}^{>})\cos(\beta\Delta\mathscr{L}_{\vec{d}}^{>})+\frac{1-\beta^2}{\beta}\exp(\Delta\mathscr{L}_{\vec{d}}^{>})\sin(\beta\Delta\mathscr{L}_{\vec{d}}^{>})]. \quad (24b)$$

† In a recent, as yet unpublished, work on *Surface Transport in Semiconductors* by R. F. GREENE, D. R. FRANKL and J. N. ZEMEL, the contribution of the unbound carriers was taken into account and the exact Poisson potential was used in the calculations. The author thanks Dr. ZEMEL for showing him the manuscript with the numerical results.

Wherein the various symbols have the following meanings†:

$$\kappa = \frac{S}{\tau}\left(\frac{m^*}{2kT}\right)^{1/2}, \tag{25a}$$

$$\epsilon = \tilde{\epsilon}/kT, \tag{25b}$$

$$\Delta\mathscr{L}_d^> \equiv \mathscr{L}_0^> - \mathscr{L}_d^> \tag{25c}$$

and

$$I\gamma = \int_{\epsilon_0}^{\epsilon_d} d\epsilon\epsilon^{-\epsilon}\exp\left[\mathscr{L}_0^>(\gamma+i\beta)\right], \tag{25d}$$

γ may be 0, 1 or 2. The subscripts $\mathscr{R}$ or $\mathscr{I}$ in the last integral refer to the real and imaginary components, respectively.

In the limit of small (or zero) magnetic fields, such that $\beta^2 \ll 1$, the scattering correction factors assume the following simple forms which are independent of β:

In the limit of small magnetic fields

$$A_r^0 = \frac{1}{\kappa_r 2\sqrt{\pi}}\int_{\epsilon_0}^{\infty} d\epsilon\epsilon^{-\epsilon}[1-\exp(\Delta\mathscr{L}_d^>)], \tag{28a}$$

$$B_r^0 = \frac{1}{\kappa_r 2\sqrt{\pi}}\int_{\epsilon_0}^{\infty} d\epsilon\epsilon^{-\epsilon}[2+(\Delta\mathscr{L}_d^>-2)\exp(\Delta\mathscr{L}_d^>)]. \tag{28b}$$

The thin film. The sample is thin enough so that both $f_1^>$ and $f_1^<$ are determined by the boundary scattering at the $z = 0$ and the $z = d$ surfaces, respectively. $f_1^>$ is given by equation (12a) and $f_1^<$ is given by equation (12b).

The average current densities and the various galvanomagnetic parameters have the same forms as those in equations (17)–(21), only now $\bar{\sigma}$, $\bar{n}$, and $\bar{p}$ are simply the bulk conductivities and carrier densities, and the integral S becomes simply d—the sample's thickness.

$$A_a^0 = \frac{1}{\kappa_a 2\sqrt{\pi}}\left\{\int_{\epsilon_0}^{\epsilon_d} d\epsilon\epsilon^{-\epsilon}[1-\exp(2\mathscr{L}_0^>)] + \int_{\epsilon_d}^{\infty} d\epsilon\epsilon^{-\epsilon}[1-\exp(\Delta\mathscr{L}_d^>)]\right\}, \tag{26a}$$

$$B_a^0 = \frac{1}{\kappa_a 2\sqrt{\pi}}\left\{\int_{\epsilon_0}^{\epsilon_d} d\epsilon\epsilon^{-\epsilon}[3+(\mathscr{L}_0^>-2)\exp(\mathscr{L}_0^>)+(\mathscr{L}_0^>-1)\exp(2\mathscr{L}_0^>)] + \right.$$

$$\left. + \int_{\epsilon_d}^{\infty} d\epsilon\epsilon^{-\epsilon}[2+(\Delta\mathscr{L}_d^>-2)\exp(\Delta\mathscr{L}_d^>)]\right\}. \tag{26b}$$

Repulsive potential. In this case $q\psi > 0$ and the scattering correction factors assume the following forms‡:

$$A_r = \frac{1}{\kappa_r 2\pi^{1/2}(1+\beta^2)}\int_{\epsilon_0}^{\infty} d\epsilon\epsilon^{-\epsilon}[(\beta^2-1)\exp(\Delta\mathscr{L}_d^>)\cos(\beta\Delta\mathscr{L}_d^>)-2\beta\exp(\Delta\mathscr{L}_d^>)\sin(\beta\Delta\mathscr{L}_d^>)+1-\beta^2], \tag{27a}$$

$$B_r = \frac{1}{\kappa_r 2\pi^{1/2}(1+\beta^2)}\int_{\epsilon_0}^{\infty} d\epsilon\epsilon^{-\epsilon}[2-2\exp(\Delta\mathscr{L}_d^>)\cos(\beta\Delta\mathscr{L}_d^>)+\frac{1-\beta^2}{\beta}\exp(\Delta\mathscr{L}_d^>)\sin(\beta\Delta\mathscr{L}_d^>)]. \tag{27b}$$

† The equations (23) and (24), as written, are pertinent to the space-charge region of thickness d adjacent to the sample's $z = 0$ surface and far removed from the other surface; in this case $\epsilon_d = 0$. These equations can pertain to a sample of arbitrary thickness (provided that the direction of the space-charge electrostatic field does not alter) in the case when ϵ_d may not be zero if a factor of 2 is introduced in front of $A_{a(\text{unbound})}$ and $B_{a(\text{unbound})}$. Thus the effects of both surfaces are included, and equations (23) and (24) describe the correction factors for a sample of arbitrary thickness d.

‡ The equations (26), (27) pertain to the case when the surface space-charge layer of thickness d is far removed from the sample's other surface. In the case that d is the sample's full thickness (and ϵ_d may not be zero), then the expressions for A and B for the repulsive potential will have the same form as they do for the attractive potential —because carriers which are in an attractive potential at the $z=0$ surface are in a repulsive potential at the $z=d$ surface.

The scattering correction factors A and B are now†:

$$A_F = \frac{2}{\kappa\pi^{1/2}(1+\beta^2)}\left\{-(1-\beta^2)F_{\mathscr{R}}+2\beta F_{\mathscr{I}}+\frac{1-\beta^2}{2}\right\},$$
$$B_F = \frac{2}{\kappa\pi^{1/2}(1+\beta^2)}\left\{-2F_{\mathscr{R}}+\frac{\beta^2-1}{\beta}F_{\mathscr{I}}+1\right\}, \tag{29}$$

where
$$F = \frac{m^*}{2kT}\int_0^\infty dv_z v_z \exp(-m^* v_z^2/2kT)\exp[-d(1-i\beta)/\tau v_z]. \tag{30}$$

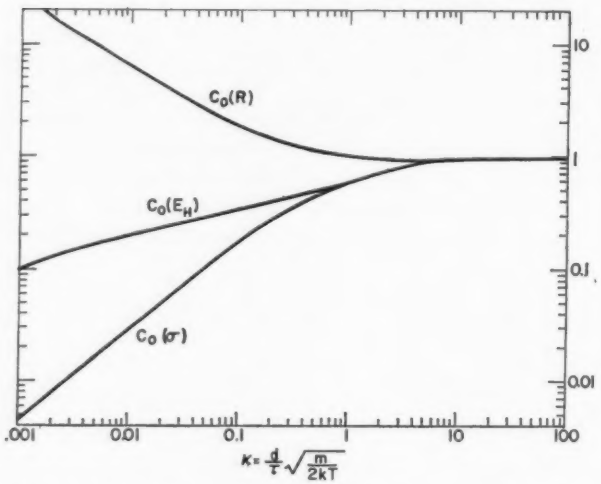


FIG. 1. Thin film, single type carrier. The relative Hall constant, Hall field (or Hall mobility, $R\sigma$) and conductivity mobility in the limit of small magnetic fields.

Subscripts $\mathscr{R}$ and $\mathscr{I}$ refer to the real and imaginary parts, respectively, and κ is a dimensionless parameter related to the ratio of the film's thickness to the mean free path (see equation (25a)): $\kappa = (m^*/2kT)^{1/2}d/\tau$. In the limit of infinitely large magnetic fields both $F_{\mathscr{R}}$ and $F_{\mathscr{I}}$ go to zero. Consequently:

$$A_F^\infty = -\frac{1}{\kappa\sqrt{\pi}}, \qquad B_F^\infty = 0.$$

In the limit of small magnetic fields:

$$A_F^0 = \frac{2}{\kappa\sqrt{\pi}}(\tfrac{1}{2}-F_{\mathscr{R}}^0), \qquad B_F^0 = 2A_F^0 -$$
$$-\left(\frac{2m^*}{\pi kT}\right)^{1/2}\int_0^\infty dv_z \exp(-m^*/2kTv_z^2)\times$$
$$\times\exp(-d/\tau v_z).$$

Figure 1 depicts the ratios of the various galvanomagnetic properties of the thin film to those of the bulk in the limit of small magnetic fields (for a single type of carrier). The following notation is used:

$$C(\sigma) = C(\overline{J_x}) = \frac{\overline{J_x}(\text{film})}{\overline{J_x}(\text{bulk})} = \frac{\sigma(\text{film})}{\sigma(\text{bulk})} = \frac{\mu(\text{film})}{\mu(\text{bulk})},$$
$$C(R) = \frac{R(\text{film})}{R(\text{bulk})},$$
$$C(E_H) = \frac{E_y(\text{film})}{E_y(\text{bulk})} = \frac{\mu^H(\text{film})}{\mu^H(\text{bulk})}, \quad \text{where } \mu^H = R\sigma.$$

† In case we are interested in the properties of a thin layer of thickness d adjacent to the $z = 0$ surface of a thick sample, then the pertinent correction factors A and B are simply half of those in equation (29).

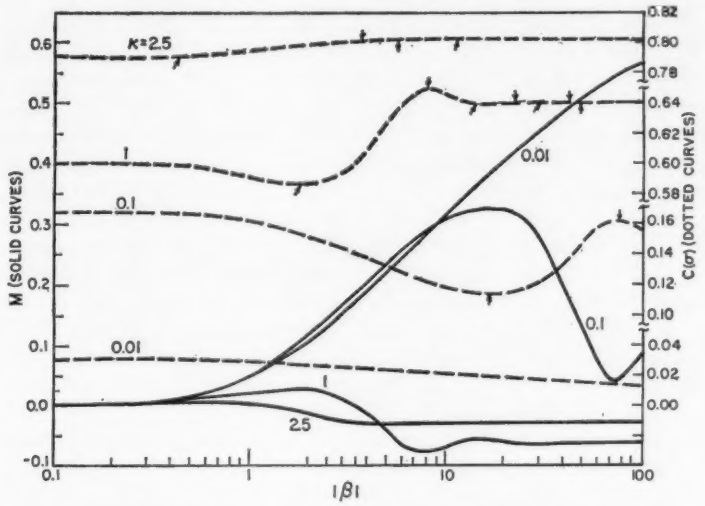


FIG. 2. Thin film, single type carrier. The relative conductivity mobility (right ordinate; the upward and downward arrows indicate positions of minima and maxima, respectively), and the magnetoresistance (left ordinate).

Note that $C(R)$ is greater than unity, and for this reason the Hall mobility of the film differs from that of the bulk to a lesser extent than do the conductivity mobilities.

Figures 2 and 3 depict the other galvanomagnetic quantities. Note that the current exhibits damped oscillations as a function of the magnetic field. The physical reason for the existence of a magnetoresistance has been mentioned in the introduction; the source of the oscillation in the conductivity is the oscillation in the speed of the carrier in a trochoidal path in perpendicular electric and magnetic fields, as was shown by CHAMBERS[11] who referred to the form of the distribution function

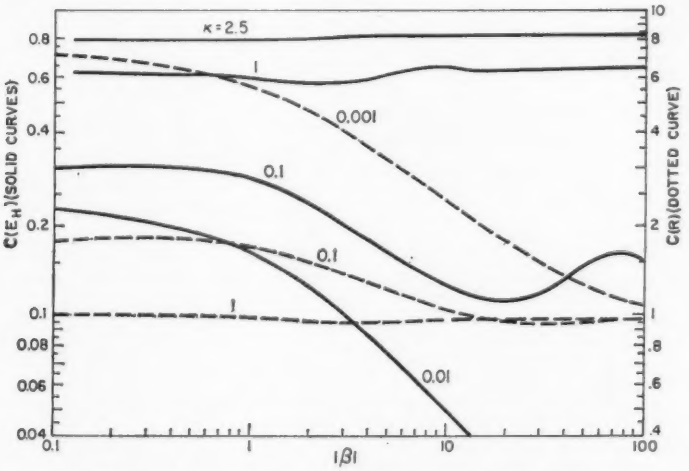


FIG. 3. Thin film, single type carrier. The relative Hall constant (right ordinate), and the relative Hall field (left ordinate; the latter equals the relative Hall mobility $R\sigma$).

in that case. A somewhat simpler point of view is the following: but for the scattering (bulk as well as surface), the carrier would move in a trochoidal path with a net velocity along the y axis only. Since the thin film's dimension d is comparable to (or smaller than) the carrier's bulk mean free path, we may concern ouselves with the surface scattering alone. The period of the orbit is $2\pi m^*c/eH$, and the time which it takes for the carriers to get from one scattering surface to the other is $d/\bar{v}_z$, where $\bar{v}_z$ is its mean thermal velocity along the thin axis. In that time, then, the carrier traverses a fraction $deH/2\pi m^*c\bar{v}_z$ of a turn. The period of oscillation will be related to a change in that fraction by unity—which means a change in the magnetic field by the amount $2\pi m^*c\bar{v}_z/ed$. Since the radius of the orbit is inversely proportional to H, it would require a larger magnetic field to impart a circular character to the orbit in a thinner film. Further, after H has become sufficiently large, the orbit is so small that a number of turns is completed before the carrier reaches the surface, and the magnetoresistance tends to a limit which is smaller than its value at the very low magnetic field.

When two kinds of carriers with different mobilities are taken into account (as in the case of intrinsic germanium), the oscillations become almost completely smeared out.

The linear truncated potential

In the computations which involve the surface space-charge layer,† we shall use a linear potential (i.e. a constant surface field) which is truncated when it reaches the value of the potential in the bulk.‡ We can expect that the main contribution of the surface scattering will take place within a mean free path or so from the surface. We may suppose, then, that this linear approximation would become worse in materials wherein the extent of the space-charge region is of the order of the mean free path or less. In such cases, especially when accompanied by very high surface fields, there are also a number of other considerations[14,15] which become of increasing concern.§ We shall take the potential to be positive; the electrostatic field, then, is negative and attractive to electrons. A_a and B_a will be the scattering correction factors for electrons; A_r and B_r will pertain to holes.

We introduce the following symbols[16,17]:

$$u = e\psi/kT, \quad L_D^2 = \frac{\epsilon kT}{4\pi e^2 n_i} \text{ and } \lambda^2 = p_0/n_0, \tag{31}$$

where ϵ is the dielectric constant of the material, and n_i is the intrinsic carrier density.

Using the truncated linear potential we obtain:

$$\begin{aligned} &S_e = (1-e^{u_s})L_D/E_s > 0, \quad S_h = (e^{-u_s}-1)L_D/E_s > 0, \\ &\mathscr{L}^{>}_{0e} = -2\alpha(\epsilon+u_s)^{1/2}, \quad \Delta\mathscr{L}^{>}_{d\,e/h} = -2\alpha[\pm(\epsilon\pm u_s)^{1/2}\mp\epsilon^{1/2}], \\ &I_\gamma = e^{u_s}-\exp[-2\alpha(\gamma+i\beta)\sqrt{u_s}]-\pi^{1/2}\alpha(\gamma+i\beta)\exp[u_s+\alpha^2(\gamma+i\beta)^2]\{erf\langle\sqrt{u_s}+\alpha(\gamma+i\beta)\rangle \\ &\qquad -erf\langle\alpha(\gamma+i\beta)\rangle\}, \end{aligned} \tag{32}$$

where

$$\alpha = -\frac{L_D}{\tau(2kT/m^*)^{1/2}E_s} > 0$$

and E_s is the electrostatic field at the surface.‖

In the calculation in this part it was assumed that electrons and holes have the same effective masses and the same relaxation time.¶ Consequently,

† The S integrals, when they weigh the carrier densities, as in equation (18), can be evaluated easily by using the correct Poisson equation for the potential. However, when these integrals appear in the expressions for A and B, they have to be evaluated using the truncated potential approximation.

‡ Such a truncated linear potential was also used, independently, by ZEMEL[12] and FLIETNER[13].

§ The present work assumes that the energy contours are spherical and that the effective mass is a scalar. Since this is not so for real semiconductors, it can be expected that comparisons of transport studies on surfaces with different crystallographic orientations would be interesting indeed.

‖ The dependence of E upon u and λ is given in Refs. 16 and 17.

¶ $m_e^*/m_0 = 0{\cdot}55$, $\mu_e = 3900\ \text{cm}^2\text{V}^{-1}\ \text{sec}^{-1}$, $L_D = 9{\cdot}76\times10^{-5}$ cm.

VOL. 14 1960

all data for holes are pertinent to electrons in a repulsive potential (the sign of u is reversed and becomes negative) after λ has been replaced by λ^{-1}. Although the calculations are over a large range of surface potentials, it has to be borne in mind that some of them are of no practical interest in germanium because they correspond to highly degenerate surface layers. The maximum values of the surface potentials allowable without causing degeneracy are:

λ^2:	10^{-7}	10^{-3}	10^{-1}	1	10	10^3	10^7
approx. max. u_s:	7	12	13	14	16	18	26

Galvanomagnetic parameters and effective mobilities

Unlike the case of the thin film, where the definition of a conductivity effective mobility or a Hall effective mobility is simple and straightforward, there are some alternate ways of choosing a meaningful definition for the effective mobilities of carriers in a surface space-charge layer. This is because of the ambiguity inherent in separating the sample's thickness into a surface layer and a bulk especially since the separation of the two* depends upon the value of the surface potential. When such a separation is accomplished, then there are some simple circuit relations[7] between the properties of the whole sample and those of the bulk and surface layers.†

Low magnetic fields

If we follow the manner in which effective mobilities have been defined in the case of the thin film, from the average current density flowing in the space-charge layer, then for electrons:

$$\mu_e^\sigma = \mu_e(1-A_e^0),$$

and for holes:

$$\mu_h^\sigma = \mu_h(1-A_h^0). \tag{33}$$

This quantity is plotted for electrons in Fig. 4. In Fig. 5 is plotted μ_e^σ at low surface potentials. Note that the effects of surface scattering on holes increase as u_s decreases, whereas the effects on electrons increase as u_s increases. The reason for this is that as u_s decreases, more holes reach the surface and they move in a layer of decreasing thickness. In the case of electrons, on the other hand, the scattering effects are due primarily to the bound carriers which move in a potential well; this well becomes deeper and narrower at the bottom as u_s increases. When u_s goes to zero, the effective mobility of the electrons still differs from their bulk mobility because of the increasing contribution of the unbound electrons in the surface layer of decreasing thickness. In a strict sense there is a discontinuity in the effective mobilities at $u_s = 0$. As u_s approaches zero, the width of the space-charge layer eventually collapses to zero. The unbound electrons, which are

* Whichever way we define the thickness of the surface space-charge layer, it depends upon the bulk properties and upon the value of the surface potential[18].

† The relationships given in part I.A of Ref. 7 are valid only for small magnetic fields. More general relationships can be derived in a straightforward manner:

$$I_x(\text{total}) = \frac{E_x}{1+\beta^2}\left\{\bar{\sigma}(1-A)d_s+\sigma d_b+\beta^2\frac{[\bar{\sigma}(1-B)d_s+\sigma d_b]^2}{\bar{\sigma}(1-A)d_s+\sigma d_b}\right\},$$

$$\sigma(\text{total}) = \frac{1}{(d_s+d_b)(1+\beta^2)}\frac{[\bar{\sigma}(1-A)d_s+\sigma d_b]^2+\beta^2[\bar{\sigma}(1-B)d_s+\sigma d_b]^2}{\bar{\sigma}(1-A)d_s+\sigma d_b},$$

$$R = \frac{\beta(1+\beta^2)(d_s+d_b)}{H}\frac{\bar{\sigma}(1-B)d_s+\sigma d_b}{[\bar{\sigma}(1-A)d_s+\sigma d_b]^2+\beta^2[\bar{\sigma}(1-B)d_s+\sigma d_b]^2}$$

where d_s and d_b are the thicknesses of the surface layer and the bulk, respectively. $\bar{\sigma}$ is given in equation (18). These reduce to expressions (2) and (3) of Ref. 7 when β^2 is neglected. What is designated there as σ_s corresponds (for low magnetic fields) to our $\bar{\sigma}(1-A^0)$.

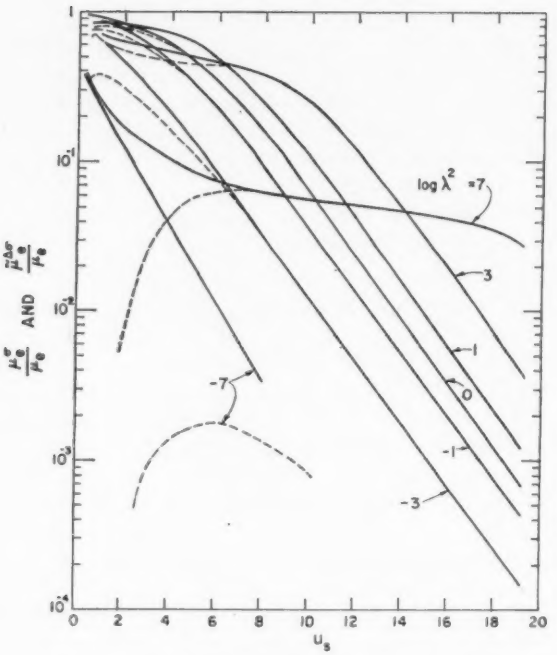


FIG. 4. μ_e^σ/μ_e (solid curves) and $\tilde{\mu}_e^{\Delta\sigma}/\mu_e$ (dashed curves).

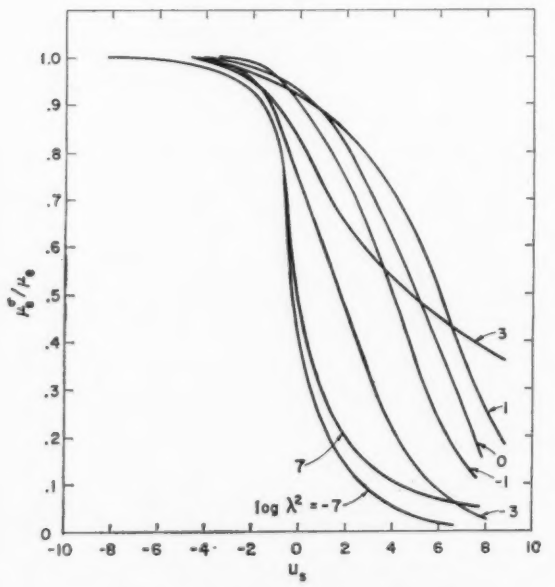


FIG. 5. μ_e^σ/μ_e for low surface potentials.

VOL. 14 1960

the only ones left, have zero mobilities in this collapsed space-charge region. Yet in a physical sense this is the case where there is no surface potential and no distinguishable surface layer. The A^0 and B^0 factors are no longer pertinent and all that need be considered is a thick sample in which the carriers have the bulk mobility. In drawing the smoothed curves this discontinuity is not brought forth.

As one alternative to the definitions of equation (33), the effective mobilities may be assigned only where

$$\tilde{\mu}_e^{\Delta\sigma} = \mu_e \left\{1 - \frac{d^{-1}S_e A_e^0}{\Delta n/n_0}\right\}, \tag{36a}$$

$$\tilde{\mu}_h^{\Delta\sigma} = \mu_h \left\{1 - \frac{d^{-1}S_h A_h^0}{\Delta p/p_0}\right\}. \tag{36b}$$

The effective mobilities $\tilde{\mu}_e^{\Delta\sigma}$ are compared with μ_e^σ in Fig. 4. $\tilde{\mu}_e^{\Delta\sigma}$ are shown in detail in Fig. 6. At high values of u_s, μ^σ and $\mu^{\Delta\sigma}$ are the same. At low values of u_s, $\tilde{\mu}_e^{\Delta\sigma} < \mu_e^\sigma$ and $\tilde{\mu}_h^{\Delta\sigma} > \mu_h^\sigma$ for reasons which are obvious from comparison of

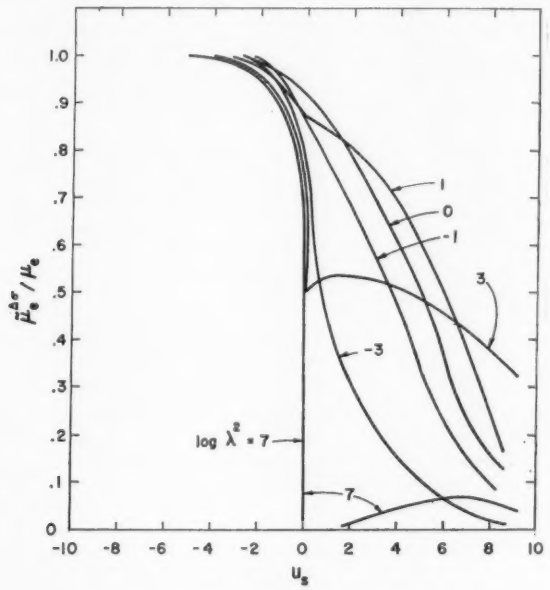


FIG. 6. $\tilde{\mu}_e^{\Delta\sigma}/\mu_e$ for low surface potentials.

to those carriers which are attracted to the surface (or those left there) by the surface potential. The conductivity in the presence of the surface potential is:

$$\sigma_1 = p_0 e\mu_h d^{-1}S_h(1-A_h^0) + n_0 e\mu_e d^{-1}S_e(1-A_e^0). \tag{34}$$

If one subtracts from this the bulk conductivity, then one may write:

$$\Delta\sigma = \Delta p e\tilde{\mu}_h^{\Delta\sigma} + \Delta n e\tilde{\mu}_e^{\Delta\sigma}, \tag{35}$$

equations (36) and (33). The effects of the scattering are ascribed to all the carriers in equation (33), whereas they are ascribed to the excess carriers alone in equation (36); this results in the inequalities noted above, because in the case of holes the excess is a negative quantity.

Instead of subtracting from σ_1 (equation 34) the bulk conductivity, we could subtract from it the conductivity of a layer of the same thickness d in which there is no surface potential. The conductivity in a layer of thickness d, adjacent to $z = 0$, in the absence of a surface potential

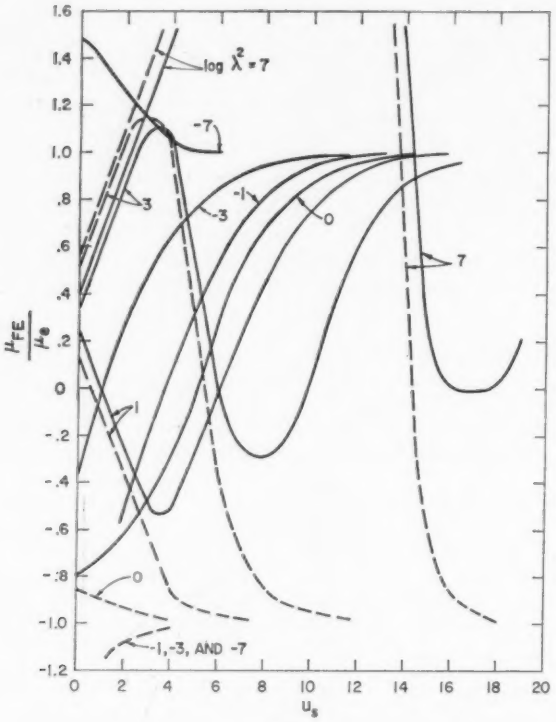


FIG. 7. The field effect mobility as compared to bulk mobility of electrons. The dashed curves show the same quantity in the absence of surface scattering (here was employed the ratio $\mu_h/\mu_e = 19/13$).

would be

$$\sigma_2 = p_0 e\mu_h(1-A^0_{Lh})+n_0 e\mu_e(1-A^0_{Le}), \quad (37)$$

where

$$A^0_L = \frac{1}{\kappa\sqrt{\pi}}(\tfrac{1}{2}-F^0_{\mathcal{R}});$$

see equation (29) and footnote!

Subtraction of equation (37) from equation (34) yields a change in conductivity which is identical in form with equation (35) but where the effective mobilities are defined differently. To obtain these new effective mobilities $\mu_e^{\Delta\sigma}$ and $\mu_h^{\Delta\sigma}$, we have to subtract from the numerators of equation (36), the quantities A^0_{Le} and A^0_{Lh}, respectively. In the effective mobilities thus defined, one always subtracts those scattering effects which are present, without the existence of a surface potential, in a layer which has the thickness of the space-charge layer. This means that for each value of λ and u_s, the thickness d has to be deduced from the spatial dependence of u, and the scattering correction factor A^0_L has to be evaluated for a layer of that thickness $d(u_s, \lambda)$. Such a procedure results in a formal separation of the scattering effects which are due to the existence of a surface potential, from those which are present in the absence of the latter. Note that $\mu_e^{\Delta\sigma}/\mu_e$ and $\mu_h^{\Delta\sigma}/\mu_h$ would both approach 1 smoothly as u_s goes to zero.

A "field effect mobility" can be defined as the rate of change of $\Delta\sigma$ equation (35) with change in the charge density in the space-charge layer; since

$$\rho = ed^{-1}(p_0 S_h - n_0 S_e)$$

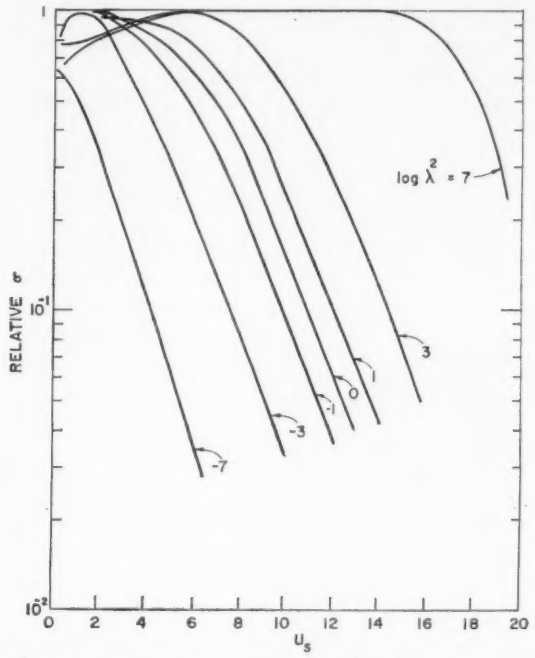


FIG. 8. The relative conductivity in the space-charge layer, i.e. the conductivity under conditions of totally diffuse scattering as compared to the conductivity without any surface scattering.

then

$$\tilde{\mu}_{\mathcal{F}.\mathcal{E}.} = \frac{\partial \Delta\sigma}{\partial \rho} = \left\{ \mu_h \lambda^2 \frac{\partial}{\partial u_s} \left[d\frac{\Delta p}{p_0} - S_h A_h^0 \right] + \right.$$
$$\left. + \mu_e \frac{\partial}{\partial u_s} \left[d\frac{\Delta n}{n_0} - S_e A_e^0 \right] \right\} \Big/$$
$$\Big/ \left\{ \lambda^2 \frac{\partial}{\partial u_s} [S_h - S_e] \right\}. \qquad (38)$$

$\tilde{\mu}_{\mathcal{F}.\mathcal{E}.}$ is plotted in Fig. 7.

If we wish to subtract out the scattering effects which obtain in the layer of thickness d in the absence of a surface potential, then we have to subtract dA_L^0 from the SA^0 terms, for the appropriate type of carriers. Note that the dependence of the thickness d upon u_s was *not* taken into account here. This definition of the field effect mobility differs from SCHRIEFFER's definition,[4] which is $\partial\Delta\sigma/\partial Q$, where Q is the total charge in the space-charge region. The latter definition has advantages with respect to correlation with field effect experiments, because Q is simply related to E_s by GAUSS's law. However, the effects of surface scattering are evident more sharply in the present $\tilde{\mu}_{\mathcal{F}.\mathcal{E}.}$.

One manner to define the Hall mobility is[7] by multiplying the Hall constant by the conductivity σ squared

$$\langle \mu^2 \rangle_H = (c/ne) R \sigma^2. \qquad (39)$$

The ratios of these Hall mobilities in the space-charge layer to the bulk mobilities for electrons and holes, respectively, are:

$$\langle \mu_e^2 \rangle_H^{1/2} = \mu_e (1 - B_e^0)^{1/2}, \quad \langle \mu_h^2 \rangle_H^{1/2} = \mu_h (1 - B_h^0)^{1/2}. \qquad (40)$$

The relative* Hall mobilities defined in this

* By "relative" we always mean the quantity in the presence of surface scattering as compared to what it would be if there were no diffuse surface scattering.

manner, are compared[5] with the corresponding μ^σ effective mobilities in Fig. 9.

If we are in a region where the conductivity is due to one type of carrier alone, then we may talk of another Hall mobility:

$$\mu^H = R\sigma, \tag{41}$$

whence

$$\mu_e^H = \mu_e \frac{1-B_e}{1-A_e}, \quad \mu_h^H = \mu_h \frac{1-B_h}{1-A_h}. \tag{42}$$

These quantities are identical with the relative Hall fields for the corresponding one carrier cases. In Fig. 10 is plotted the relative Hall field,

$$\frac{\lambda^2 S_h(1-B_h^0)-S_e(1-B_e^0)}{\lambda^2 S_h(1-A_h^0)+S_e(1-A_e^0)} \Big/ \frac{\lambda^2 S_h - S_e}{\lambda^2 S_h + S_e}, \tag{43}$$

and the regions where the single carrier model applies are designated by heavy lines. Note that these effective Hall mobilities, which are proportional to $(1-B)/(1-A)$, are sometimes larger than the conductivity effective mobilities μ^σ, which are proportional to $1-A$, and sometimes smaller; in the thin film case they were usually larger.

In Fig. 11 are plotted the relative Hall constants

$$\frac{\lambda^2 S_h(1-B_h^0)-S_e(1-B_e^0)}{[\lambda^2 S_h(1-A_h^0)+S_e(1-A_e^0)]^2} \Big/ \frac{\lambda^2 S_h - S_e}{[\lambda^2 S_h + S_e]^2}, \tag{44}$$

and the regions where this quantity corresponds to the case of a single type of carrier are marked.

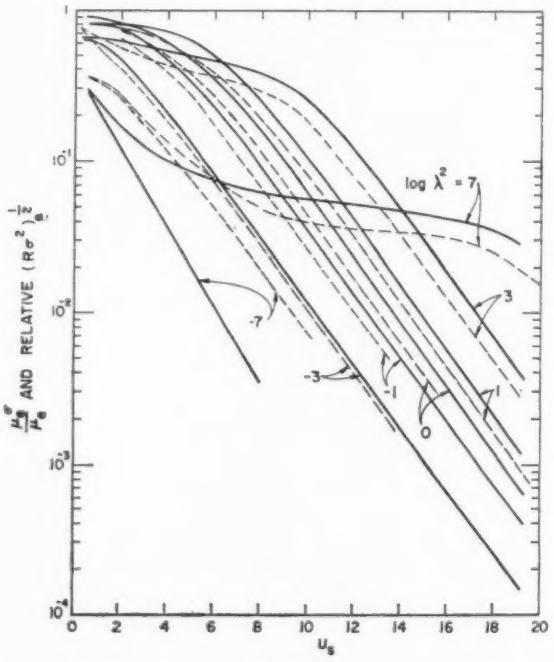


FIG. 9. The relative Hall mobility ($\langle R\sigma^2\rangle^{1/2}$) for electrons (dashed curve) compared to μ_e^σ/μ_e (solid curve).

In Figs. 10 and 11 it is noted that the curves for $\lambda^2 = 10^7$ and $\lambda^2 = 10^3$ become negative (the negative portions are not drawn in the figures). This is due to the fact that the Hall field and the Hall constant in presence of scattering do not reverse their signs where they would do so in the absence of scattering.

Arbitrary magnetic fields

We shall restrict ourselves here to the case where the conduction process is due to a single

VOL. 14 1960

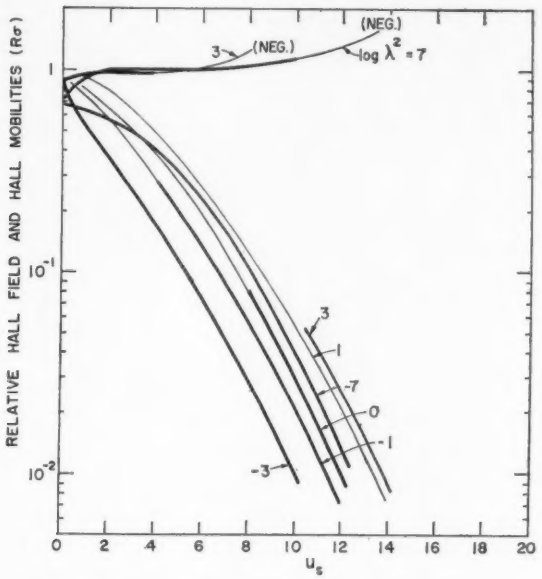


FIG. 10. The relative Hall field. The heavy lines represent μ_h^H/μ_h for $\lambda^2 = 10^7$ and 10^3 near the origin, and all other heavy lines represent μ_e^H/μ_e; these represent the ranges where the conductivities are due to a single type of carrier.

type of carrier alone.* The conductivity effective mobility becomes

$$\mu_e^\sigma = \mu_e\{1-A_e+\beta_e^2(1-B_e)^2/(1-A_e)\}/(1+\beta_e^2). \tag{45}$$

The Hall effective mobility $R\sigma$ is

$$\mu_e^H = \mu_e \frac{1-B_e}{1-A_e}. \tag{46}$$

These quantities are plotted in Fig. 12. Note the different behavior exhibited by these quantities in the present case as compared to their behavior in the case of the thin film (Figs. 2 and 3). In the case of the thin film, the oscillations in the effective mobilities (in the case of the single type of carrier) were intimately tied to the relation between the film's thickness and the radius of the carrier's orbit. The thickness of the potential well in which a carrier is constrained to move depends upon the latter's energy—hence any oscillatory behavior would be smeared out. Only in the extremely narrow channels are the effects of the magnetic field noticeable.

CONCLUSION

The galvanomagnetic properties have been formulated for thin films and for surface space-charge layers. They have been evaluated in terms of normalized parameters for the one carrier case, and for room temperature germanium in the two carrier case.† It was noticed that the effects of the surface potential are severe and there is no simple way of comparing the behavior of the various parameters to the case of the thin film. One should be aware of the limitation in the approach adopted

* This actually pertains to a sample of total thickness d with a space-charge layer; it was noted in the footnote† on page 281 that for arbitrary magnetic fields one could not separate the bulk from the surface layer as one had done for low magnetic fields.

† In Ref. 6 it was noted that there are significant contributions from the light holes. These were not taken into account here, since we assumed that mobilities of all carriers are the same.

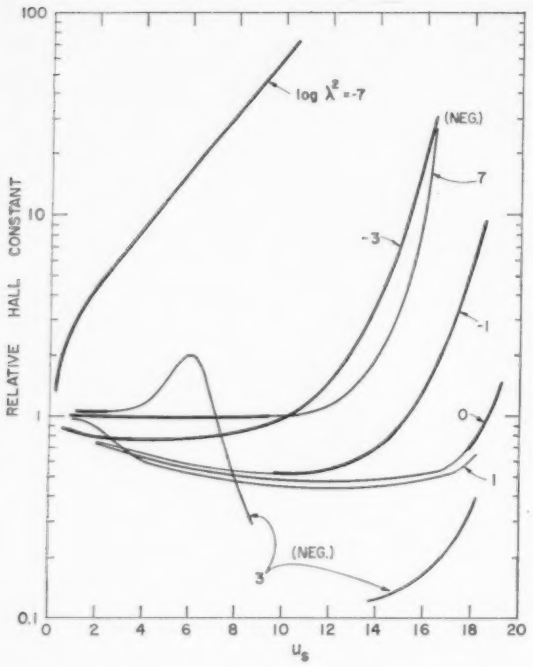


FIG. 11. The relative Hall constant. The heavy lines represent $1-B^0/(1-A^0)^2$ which is the relative Hall constant for a single type of carrier. Near the origin, for $\lambda^2 = 10^7$ and 10^3, it is the relative Hall constant for holes; elsewhere for electrons.

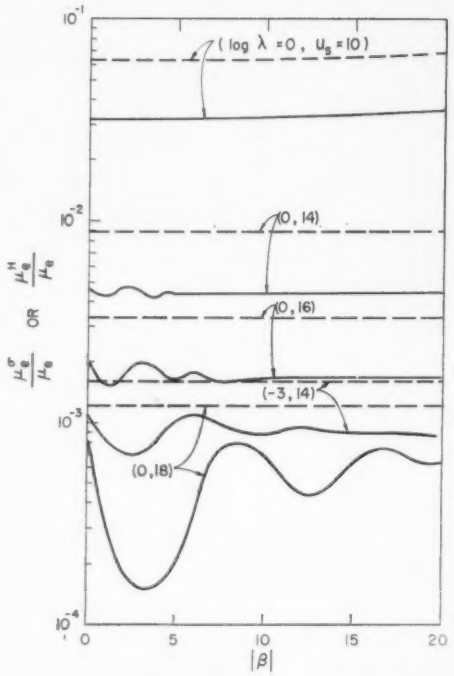


FIG. 12. μ_e^σ/μ_e (dashed curves) and $\mu_{e\mathcal{E}}^H/\mu_e$ (solid curves) for some selected values of λ^2 and u_s as functions of the magnetic field.

here, which uses the bulk relaxation time in the transport equation. As the region under consideration becomes thinner and thinner one is hard put to justify this procedure, because no part of the region can be classified as bulk. What seems to be required is a fundamental approach in terms of integral collision operators to define the pertinent relaxation time which is to be used in the transport equation. It is most unlikely that such a relaxation time for the thin film or the space-charge layer would not differ from the one in the bulk. Another significant point is that had the anisotropy of the energy surfaces been taken into account, then the effective mobilities in the surface layer would have been smaller than the bulk mobilities (in a manner dependent on crystallographic orientation) even in the case of perfectly specular reflection.

Acknowledgements—The author wishes to thank Dr. I. ADAWI for a number of helpful discussions and Dr. J. N. ZEMEL, as well as Dr. A. M. GOODMAN, for critical reading of the original manuscript. Deserving of particular gratitude for their indispensable aid in programming all the numerical work for the IBM–650 Computer are Messrs. F. HERZFELD (for the space-charge layer calculations) and A. H. SIMON (for the thin film calculations).

APPENDIX

Mixed specular and diffuse scattering

In this case a fraction r of all carriers impinging upon the surface undergo elastic collisions and are reflected specularly. Consider the situation at the $z = 0$ surface: the distribution function of carriers leaving that surface is $f_0+f_1^>(v_z)$, that of carriers approaching that surface is $f_0+f_1^<(v_z)$. A fraction of the carriers moving toward the surface are reflected specularly, after which they are described by the distribution function $r[f_0+f_1^<(-v_z)]$; the remainder of the carriers are scattered diffusely and are described by the distribution function g_1[19]. Hence

$$f_0+f_1^>(v_z) = r[f_0+f_1^<(-v_z)]+g_1.$$

At the scattering surface $z = 0$, the function g_1 which

describes the diffusely scattered carriers has to be independent of their sense of motion. Consequently:

at $z = 0$

$$f_1^{>}(v_z) = rf_1^{<}(-v_z). \tag{A 1a}$$

The condition on $f_1^{<}(v_z)$ depends upon the specific case under study, as shown below.

Attractive potential. In addition to the requirement (A 1a) at $z = 0$, there is the condition of specular reflection at the potential incline:

at $z = z_p$

$$f_1^{>}(v_z) = f_1^{<}(-v_z). \tag{A 1b}$$

The distribution functions which satisfy these requirements are:

$$f_1^{>}(v_z) = f_1^{0\beta} - \alpha^{>}(v_z)\frac{q\tau v f_0}{kT(1+\beta^2)} \times$$

$$\times \exp(\Delta\mathscr{L}^{>})\{X\cos(\beta\Delta\mathscr{L}^{>}) + \Upsilon\sin(\beta\Delta\mathscr{L}^{>})\}, \tag{A 2a}$$

and

$$f_1^{<}(-v_z) = f_1^{0\beta}[1+\alpha^{<}(-v_z)e^{\mathscr{L}}] - \alpha^{<}(-v_z)\frac{q\tau v f_0}{kT(1+\beta^2)}e^{\mathscr{L}} \times$$

$$\times\{X[\exp(\mathscr{L}_0^{>})\cos(\beta\mathscr{L}_0^{>}) + \cos(\beta\mathscr{L})] + \Upsilon[\exp(\mathscr{L}_0^{>})\sin(\beta\mathscr{L}_0^{>}) + \sin(\beta\mathscr{L})]\}, \tag{A 2b}$$

where:

$$\alpha^{\gtrless}(v_z) = \frac{1-r}{1+r\exp(\pm\mathscr{L}_0^{>})\{1-[1+\exp(\pm\mathscr{L}_0^{>})][\cos(\beta\mathscr{L}_0^{>}) \pm \Upsilon X^{-1}\sin(\beta\mathscr{L}_0^{>})]\}} \tag{A3}$$

and

$$\alpha^{<}(-v_z) = \alpha^{>}(v_z).$$

When the scattering is totally diffuse, $r = 0$ and the α coefficients are all unity. The distribution functions given by equations (A 2) then become those derived previously and given by equations (7) and (11), respectively. When the scattering is purely specular, $r = 1$ and the α coefficients are all zero. The distribution functions then become simply $f_1^{0\beta}$, that is the bulk functions given by equation (9).

When the sample is thick, $f_1^{<}$ for the unbound carriers is simply $f_1^{0\beta}$ and $\alpha^{>}$ reduces to $1-r$.

Repulsive potential. The functions $f_1^{>}$ and $f_1^{<}$ for the unbound carriers are identical in form with the analogous functions in the case of the attractive potential. $f_1^{>}$ for the bound carriers can be obtained in a fashion similar to that employed in obtaining $f_1^{<}$ for the bound carriers in the case of an attractive potential. As was stated before, however, $f_1^{<}$ will be simply $f_1^{0\beta}$ because of the sample's thickness, and $\alpha^{>}$ will be simply $1-r$.

No surface potential. In this case, which is restricted to thin films, there is a condition at the $z = d$ surface in addition to equation (A 1a):

at $z = d\delta$.

$$f_1^{<}(v_z) = rf_1^{>}(-v_z).$$

The general forms of the distribution functions which satisfy these requirements are:

$$f_1^{>}(v_z) = f_1^{0\beta} - \alpha^{>}(v_z)\frac{q\tau v f_0}{kT(1+\beta^2)} \times$$

$$\times \exp(-z/\tau v_z)\left\{X\cos\frac{\beta z}{\tau v_z} - \Upsilon\sin\frac{\beta z}{\tau v_z}\right\}, \tag{A 4a}$$

$$f_1^{<}(v_z) = f_1^{0\beta} - \alpha^{<}(v_z)\frac{q\tau v f_0}{kT(1+\beta^2)} \times$$

$$\times \exp[(d-z)/\tau v_z]\left\{X\cos\beta\frac{d-z}{\tau v_z} + \Upsilon\sin\beta\frac{d-z}{\tau v_z}\right\}, \tag{A 4b}$$

where:

$$\alpha^{\gtrless}(v_z) = \frac{1-r}{1-r\exp(\mp d/\tau v_z)\left\{\cos\frac{\beta d}{\tau v_z} \mp \Upsilon X^{-1}\sin\frac{\beta d}{\tau v_z}\right\}}.$$

REFERENCES

1. Fuchs K., *Proc. Camb. Phil. Soc.* **34,** 100 (1938).
2. Sondheimer E. H., *Phys. Rev.* **80,** 401 (1950).
3. Babiskin J. and Siebenmann P. G., *Phys. Rev.* **107,** 1249 (1957).
4. Schrieffer J. R., *Phys. Rev.* **97,** 641 (1955).
5. Zemel J. N., *Phys. Rev.* **112,** 762 (1958).
6. Zemel J. N. and Petritz R. L., *Phys. Rev.* **110,** 1263 (1958).
7. Petritz R. L., *Phys. Rev.* **110,** 1254 (1958).
8. Montgomery H. C. and Brown W. L., *Phys. Rev.* **103,** 865 (1956).
9. Wilson A. H., *The Theory of Metals.* Cambridge University Press, Cambridge (1954).
10. Sneddon I. N., *Elements of Partial Differential Equations.* McGraw-Hill, New York (1957).
11. Chambers R. G., *Proc. Roy. Soc.* **A202,** 378 (1950).
12. Zemel J. N., *Bull. Amer. Phys. Soc.* (ser. 2), **3,** 255 (1958).

13. FLIETNER H., *Ann. Phys., Lpz.* **3,** 396 (1956).
14. HAM F. S. and MATTIS D., *Phys. Rev.* **99,** 1656 (1955).
15. SCHRIEFFER J. R., *Semiconductor Surface Physics,* (edited by KINGSTON R. H.) p. 55. University of Pennsylvania Press, Philadelphia (1957).
16. KINGSTON R. H. and NEUSTADTER S. F., *J. Appl. Phys.* **26,** 718 (1955).
17. GARRETT C. G. B. and BRATTAIN W. H., *Phys. Rev.* **99,** 376 (1955).
18. DOUSMANIS G. C. and DUNCAN R. C. Jr., *J. Appl. Phys.* **29,** 1627 (1958).
19. DINGLE R. B., *Proc. Roy. Soc.* A**201,** 545 (1950).

J. Phys. Chem. Solids Pergamon Press 1960. Vol. 14. pp. 291–298. Printed in Great Britain.

SURFACE TRANSPORT THEORY

R. F. GREENE

United States Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland

VOL. 14 1960

1. SURFACE CONDUCTIVITY

ONE can induce in a semiconductor slab a surface space-charge layer containing enough excess free carriers to alter strongly the conductance parallel to the surface.[1] This is, of course, characteristic of semiconductors: Metals have so many bulk electrons available for electrostatic shielding that space-charge layers cannot extend much beyond an interatomic spacing into the bulk. The key quantities in surface transport are then the surface excesses[2–5] of electrons and holes:

$$\Delta N = \int_0^\infty dz(n-n_B) \tag{1a}$$

$$\Delta P = \int_0^\infty dz(p-p_B), \tag{1b}$$

(n and p are the electron and hole concentrations. Subscript B denotes bulk values) and the corresponding effective surface mobilities μ_{nS} and μ_{pS}: These are defined by writing the integrated current density across a thick slab (thickness $2d$) as

$$I = \int_0^{2d} dz\, j(z) = I_0 + 2E\Delta\sigma \tag{2a}$$

where

$$\Delta\sigma = e\mu_{nS}\Delta N + e\mu_{pS}\Delta P. \tag{2b}$$

Here I_0 is the flat-band value and $\Delta\sigma$ is the surface conductivity corresponding to carrier excesses on each face.

By means of Poisson's equation ΔN and ΔP can be expressed as a function of the surface potential excess $(u_S-u_B)(kT/e)$. It is not hard to see that $\Delta\sigma$ should have a minimum[6] $\Delta_m\sigma$ at some value of $\Delta_S u = (u_S-u_B)$ for which ΔN and ΔP are both small. In fact if μ_{nS} and μ_{pS} are taken to be constants one can express $\Delta\sigma$ as a function of $\Delta_S u$ by means of the exact formulas[4]

$$\Delta N = n_B L_D \int_0^{\Delta_S u} d\Delta u\,(\exp\Delta u - 1)/F(\Delta u, u_B), \qquad (\Delta u = u - u_B) \tag{3a}$$

$$\Delta P = p_B L_D \int_0^{\Delta_S u} d\Delta u\,(\exp(-\Delta u) - 1)/F(\Delta u, u_B) \tag{3b}$$

where[7]

$$L_D = (\kappa kT/8\pi n_i e^2)^{1/2} \tag{3c}$$

and

$$F = 2\sinh(\tfrac{1}{2}\Delta u)[\cosh u_B + g(\Delta u)\sinh u_B]^{1/2} \tag{3d}$$

and

$$g = (\sinh\Delta u - \Delta u)/(\cosh\Delta u - 1) \tag{3e}$$

Then the conductivity minimum occurs at[3]

$$\Delta_{Sm}u = \ln(p_B\mu_{nS}/n_B\mu_{pS}) \tag{4}$$

When the bulk is intrinsic an explicit expression for $\Delta\sigma$ results:

$$\Delta N = 2L_D n_i\,[\exp(\Delta_S u/2) - 1] \qquad (u_B = 0) \tag{5a}$$

$$\Delta P = 2L_D n_i[\exp(-\Delta_S u/2) - 1] \qquad (u_B = 0) \tag{5b}$$

One then finds

$$\Delta\sigma = 4L_D n_i e(\mu_{nS}\mu_{pS})^{1/2}[\cosh\tfrac{1}{2}(\Delta_{Sm}u - \Delta_S u) - \cosh\tfrac{1}{2}\Delta_{Sm}u], \quad u_B = 0 \tag{6}$$

Then from measurements of $(\Delta\sigma - \Delta_m\sigma)/\Delta_m\sigma$ one can get u_S, using equation (4). When $u_B \neq 0$, one must use equation (3) and proceed numerically.

This situation became less simple when SCHRIEFFER[6] pointed out that surface scattering should depress μ_{nS} and μ_{pS} below μ_n and μ_p, respectively, and that μ_{nS} and μ_{pS} should depend

upon $\Delta_S u$. As $\Delta_S u$ is increased, for example, the average excess electron is held closer to the surface and therefore suffers more surface scattering and has a lower mobility. SCHRIEFFER gave a Boltzmann equation treatment of the transport, following FUCHS'[8,9] analysis of transport in thin metal films. FUCHS represents the surface scattering by a boundary condition on the distribution function. The Fuchs b.c., corresponding to a diffusely scattering surface, is expressed as the suppression of the transport term f_1 of f for electrons *leaving* the surface:

$$f_1 = 0 \text{ for } z = 0 \text{ and } v_z \geqslant 0 \tag{7}$$

SCHRIEFFER took the key step of introducing the space-charge potential explicitly into the equilibrium distribution function f_0 and thence into the Boltzmann equation. The resulting boundary value problem was solved by an elegant transformation, and an expression for the current obtained. This gave formulas for μ_{nS}/μ_n which decrease monotonically from unity to zero as the bands are bent downward at the surface.

The exact shape of the space charge potential function has to be obtained numerically[4] from Poisson's equation. That function is generally involved in surface transport integrals and makes surface theory much more awkward mathematically than bulk theory. Extensive numerical evaluation is usually required except when the crudest assumptions and approximations are made (e.g. it has so far been necessary to treat the bulk relaxation time τ as a constant).

In SCHRIEFFER's treatment approximate potential functions are used which are accurate near the surface but which diverge linearly into the bulk instead of asymptoting to the finite value u_B. (The expression for E_z used in equation (22) of Ref. 6 becomes a non-zero constant when $e\psi/kT = \Delta u$ is set equal to $e\psi_S/kT = \Delta_S u$). There is, moreover, a corresponding deficiency in the formalism (i.e. in the boundary value problem) which suppresses the bulk current and which gives incorrect expressions for μ_{nS}/μ_n.

Several attempts have been made to improve these aspects of the theory. ZEMEL[10] used a linearized potential function, truncated at u_B, in the SCHRIEFFER formalism. FRANKL[11] and, independently, GREENE[12] used the exact potential function and revised the formalism. Recently FLIETNER[13] has used a truncated-linear potential but with an incorrect Boltzmann boundary value treatment. The theory, as reworked by GREENE *et al.*[14] may be compared with SCHRIEFFER's treatment as follows*: SCHRIEFFER introduces the space-charge potential function

$$u = \ln(n/n_i) = -\ln(p/n_i) \tag{8}$$

explicitly into the equilibrium distribution function

$$f_0 = n_i(m/2\pi kT)^{3/2} \exp(u - mv^2/2kT). \tag{9}$$

Applying this to the Boltzmann equation

$$v \cdot \nabla f - e\boldsymbol{E} \cdot \nabla_p f = f_1/\tau, \; f = f_0 + f_1, \tag{10}$$

and taking E_x = (applied field), $E_z = -(kT/e)(\partial u/\partial z)$ = (equilibrium field), applying planar symmetry conditions, the Ohmic transport equation

$$f_1 = f_{11} - \tau v_z(\partial f_1/\partial z) + (e\tau E_z/m)(\partial f_1/\partial v_z) \tag{11a}$$

was obtained,[6] where

$$f_{11} = e\tau v_x E_x f_0' = -e\tau v_x E_x f_0/kT. \tag{11b}$$

The $E_z f_1$ term of equation (11a) is missing from Ohmic transport equations for metals[8, 9] because there macroscopic equilibrium fields are screened out. SCHRIEFFER simplified equation (11) by LAGRANGE's method[15] whereby a special solution

$$\nu(z,v_z) = mv_z^2/2kT - u + \text{constant} \tag{12}$$

provides a transformation from (z, v_z) space to (ν, v_z) space in which equation (10) becomes

$$(\partial/\partial v_z) f_1 e^K = f_{11}(\partial/\partial v_z) e^K, \; (\nu = \text{const.}) \tag{13}$$

$$K = K(v_z, v_{z1}) \equiv -\int_{v_{z1}}^{v_z} m \, dv_z/e\tau E_z, \; (\nu = \text{const.}) \tag{14}$$

The Fuchs b.c. is then applied[6] in the form

$$f_1(\nu, v_{zS}) = 0 \tag{15}$$

where v_{zS} is the velocity of an electron *leaving* the surface with "energy" ν. If $u(z)$ is assumed to increase monotonically and without bound, as in SCHRIEFFER's treatment,[6] equation (13) may be integrated over the range $(v_{zS} \geqslant v_z \geqslant -v_{zS})$. But if $u(z)$ is bounded above by u_B then the foregoing statement is incorrect for electrons with $\nu > 0$,

* The notation of Ref. 14 is used throughout.

taking the constant in equation (12) as u_B. For these electrons, which are not bound to the surface space-charge layer, the positiveness of $mv_z^2/2$ splits the range of v_z into two non-contiguous parts:

$$v_{zS} \geqslant v_z \geqslant v_B \text{ and } -v_{zB} \geqslant v_z \geqslant -v_{zS}, \quad (16)$$

where v_{zB} is the positive z-velocity of an electron in the bulk having $\nu > 0$. Thus for $\nu > 0$ an additional b.c. is needed. This must be taken as[14]

$$f_1(-v_{zB}) = f_{11} \quad (\nu > 0) \quad (17)$$

in order to make f_1 fit smoothly onto the bulk distribution function.

It may be noted that one cannot specify *both* $f_1(\nu, v_{zS})$ and $f_1(\nu, -v_{zS})$ because the one may be calculated from the other via equation (13).

From the formal solution of this corrected boundary value problem a new expression for the current results,[14] having a flat-band term (containing the bulk current) plus a surface-potential dependent term. The latter provides revised expressions for μ_{nS} and μ_{pS}. For a slab (with two surfaces) of width $2d \gg L_D$

$$I = eE_x[n_B\mu_n(2d-2\lambda_n)+p_B\mu_p(2d-2\lambda_p)]+ + 2eE_x(\mu_{nS}\Delta N+\mu_{pS}\Delta P) \quad (18a)$$

where

$$\lambda_n = \tau_n(kT/2\pi m)^{1/2} \quad (18b)$$

is a thermal mean free path and where

$$\mu_{nS}/\mu_n = 1-(\lambda_n n_B/\Delta N)[\exp(\Delta_S u)-1], \quad \Delta_S u \leqslant 0 \quad (18c)$$

$$\mu_{nS}/\mu_n = 1-(\lambda_n n_B/\Delta N)\int_{-\Delta_S u}^{0} dv e^{-\nu}\{1-\exp[2K_n(v_{zS},0)]\}, \quad \Delta_S u \geqslant 0 \quad (18d)$$

and

$$2K_n(v_{zS},0) = -(L_D/\lambda_n)\pi^{-1/2}\int_{-\nu}^{\Delta_S u} d\Delta u(\Delta u+\nu)^{-1/2}/F \quad (18e)$$

with similar expressions for holes.

Numerical evaluation,[14] using digital computors, was made of the expressions (18c) and (18d) for a range of values of u_S and u_B and for various values of the parameter (L_D/λ_n) or (L_D/λ_p). corresponding to electrons and holes in germanium and silicon at 300°K. Using the data compiled by CONWELL[16] one finds

	Ge	Si	
L_D/λ_n	33·3	2,270	
L_D/λ_{ph}	54·8	4,610	heavy holes
L_D/λ_{pl}	18·3	—	light holes

Conductivity effective masses were used—somewhat artificially since a simple spherical band model was used in equation (18). The light and heavy hole mobilities in germanium were obtained from the formula

$$(p_l+p_h)\mu_p = p_h\mu_{ph}+p_l\mu_{pl}, \quad (19)$$

using $p_l/p_h = 0{\cdot}0423$ given by LAX and MAVROIDES[17] and using the ratio $\mu_{pl}/\mu_{ph} = 8$ given by WILLARDSON *et al.*[18] and by ZEMEL and PETRITZ[19]. Although the corresponding situation in silicon is less clear, p-type silicon was represented in the calculation simply by neglecting the light holes.

Results for light holes in germanium are shown in Figs. 1 and 2. Cusps occur at the flat band points where the two expressions (18c) and (18d) fit together giving[14]

$$\mu_{pS}/\mu_p = 1-(\lambda_p/L_D)(\cosh u_B)^{1/2} \quad (\Delta_S u = 0) \quad (20)$$

This can be appreciably less (e.g. $\mu_{pS}/\mu_p = 0{\cdot}72$ for $u_S = u_B = 4$ in Fig. 2) than SCHRIEFFER's result which is unity.

A truncated linear approximation of the potential function,

$$\Delta u = \Delta_S u(z_1-z)/z_1 \quad \text{for } z \leqslant z_1 \quad (21a)$$

$$\Delta u = 0 \quad \text{for } z \geqslant z_1 \quad (21b)$$

$$\Delta_S u = F(\Delta_S u, u_B)z_1/L_D, \quad (21c)$$

substituted in the expression (18d) gives

$$\mu_{nS}/\mu_n = 1-(\exp\Delta_S u-\Delta_S u-1)^{-1}\exp(\Delta_S u)\times \times \{\exp\alpha^2[\text{erf}\alpha-\text{erf}(\alpha-(\Delta_S u)^{1/2})]- -[1-\exp(-2\alpha(\Delta_S u)^{1/2})/(\alpha\pi^{1/2}\exp\Delta_S u)]\} \quad (22a)$$

where

$$\alpha = L_D/\lambda_n\pi^{1/2}F(\Delta_S u, u_B) = (2mkT)^{1/2}/e\tau E_{zS}. \quad (22b)$$

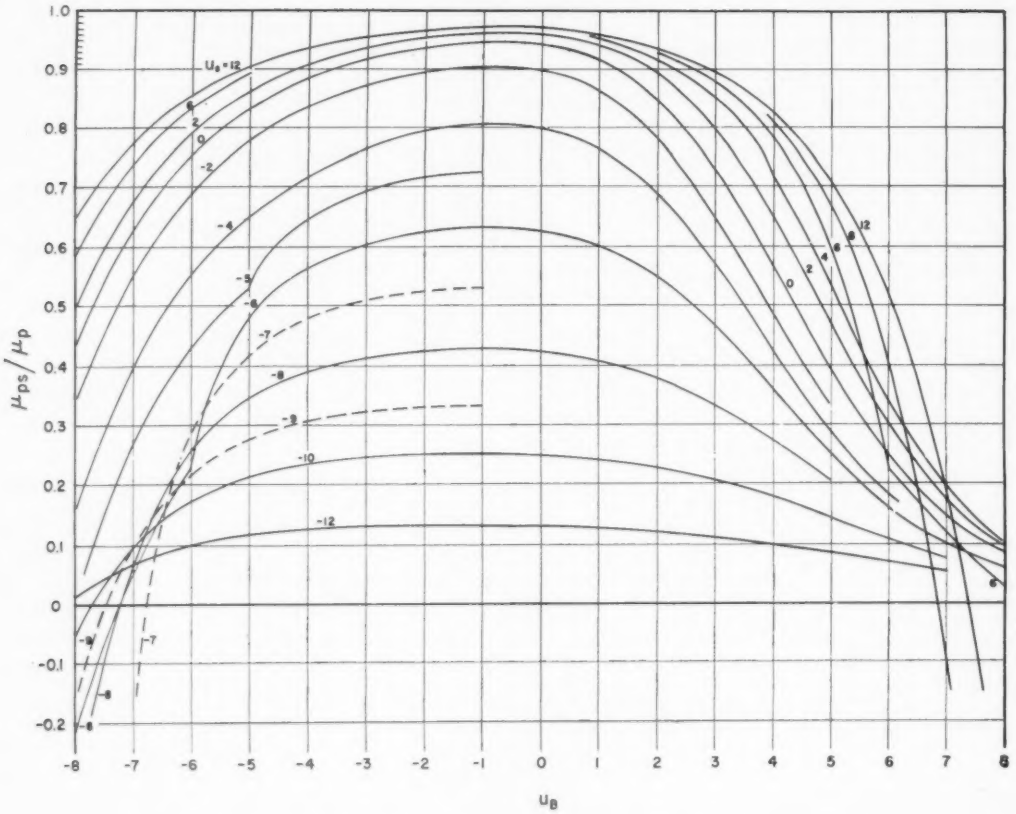


FIG. 1. Surface mobility of light holes in germanium at 300°K vs. bulk potential u_B for various values of surface potential u_S (theoretical). After GREENE *et al.*[14]

This reduces to SCHRIEFFER's linear approximation[6]

$$\mu_{nS}/\mu_n = 1-\exp(\alpha^2)(1-\operatorname{erf}\alpha) \qquad (23a)$$

when

$$\alpha(\Delta_S u)^{1/2} \ll 1 \quad (\Delta_S u \gg 1) \qquad (23b)$$

The accuracy of equation (22a) is rather low because of the crudeness of the approximation: It gives twice the exact value of the deviation $(1-\mu_{nS}/\mu_n)$ in the flat band limit. The fit of equation (22a) to the exact result (equation (18d)) could be improved by using a field in expression (22b) somewhat less than the exact surface value E_{zS}. This is expected intuitively, since the "average" excess electron is one or two kT above the channel bottom and such an electron "sees" a field of about $E_{zS}/2$.

2. QUALITATIVE ARGUMENTS

The main features of semiconductor surface transport can be inferred from rough intuitive arguments. Consider a slab of thickness $2d$ having an accumulation layer of thickness L on either face. From equation (3) it can be seen that this thickness is (when $\Delta_S u$ is small)

$$L \approx L_D(\cosh u_B)^{-1/2}. \qquad (24)$$

On the other hand, surface scattering can be thought of as producing a stagnation layer of thickness $\approx \lambda$ on the surface. In more highly doped semiconductors as $|u_B|$ increases, the ratio λ_n/L increases, a greater fraction of the excess carriers are in the stagnation layer, and μ_{nS}/μ_n decreases. This effect can be seen in Figs. 1 and 2.

VOL. 14 1960

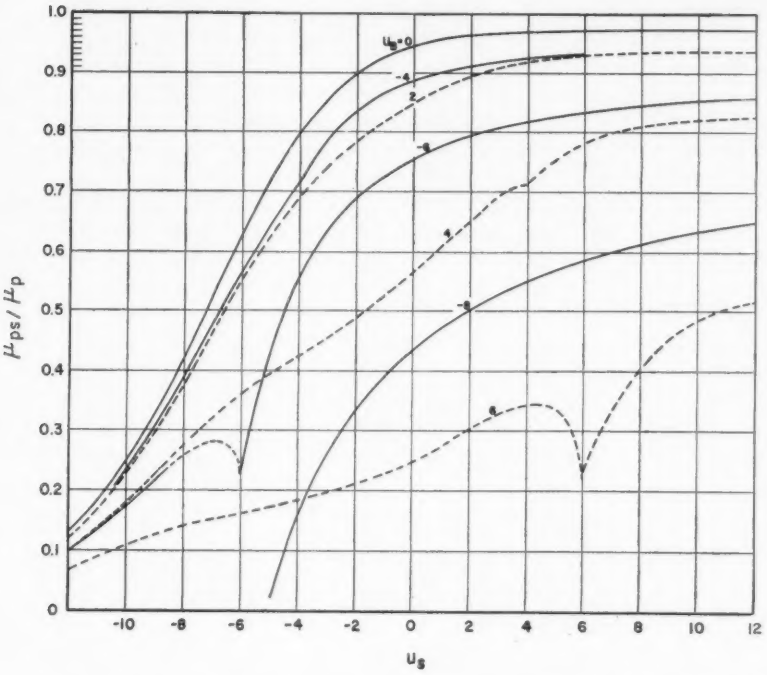


FIG. 2. Surface mobility of light holes in germanium at 300°K vs. surface potential u_S for various values of bulk potential u_B (theoretical). After GREENE *et al.*[14]

In this simple model the two accumulation layers carry a current

$$2E_x e\mu_n n_S(L-\lambda). \tag{25a}$$

and the bulk region carries the current

$$eE_x\mu_n n_B(2d-2L). \tag{25b}$$

The total current is then

$$I \approx eE_x n_B\mu_n(2d-2\lambda)+2eE_x[L(n_S-n_B)]\mu_n(1-\lambda/L) \tag{25c}$$

whence

$$\mu_{nS}/\mu_n = 1-\lambda/L. \tag{25d}$$

(The flat band result, equation (20), is now easily understood.) A correction should be made to equation (25c) to account for the fact that electrons moving away from the surface can be surface scattered (by "rebounding" from the potential wall at L) as well as those moving toward the surface. It can be shown[14] that they will then have a probability

$$(\lambda/L)\,(1-e^{-2L/\lambda}) \tag{26}$$

instead of (λ/L) of suffering a surface scattering before being bulk scattered. Inserting this correction in equation (25c) gives

$$\mu_{nS}/\mu_n \approx 1-(\lambda/L)\,(1-e^{-2L/\lambda}) \tag{27}$$

The formal resemblance to equation (18d) emerges when K is rewritten, using expression (12), as

$$K = -\int_0^{L(y)} dz/\tau v_z \approx -L/\lambda. \tag{28}$$

3. THIN SLABS: PARTIALLY SPECULAR SURFACES

Since the proportion of surface to bulk currents is increased in thinner slabs, it is useful to extend the foregoing theory to cover slabs whose thickness $2d$ may be small enough for the two surface space-charge layers to coalesce ($\sim 10^{-4}$ cm). It would also

be useful to extend the formalism so that experimental means could decide to what extent the hypothesis of diffuse surface scattering is preferable to that of specular surface reflection. These two generalizations are easily made,[14] using the appropriate solution of Poisson's equation and applying FUCHS's scheme[8, 9] of interpolating between the diffuse model, equation (15), and the specular model: $f_1(v_{zS}) = f_1(-v_{zS})$. An interpolation parameter w is introduced such that

$$f_1(v_{zS}) = wf_1(-v_{zS}),\ 0 \leqslant w \leqslant 1. \qquad (29)$$

One interprets w loosely as the probability of specular reflection rather than diffuse scattering. Solution of this boundary value problem[14] again yields the current expression (18a), but with $\lambda_n(1-w_n)$ replacing λ_n, and $\lambda_p(1-w_p)$ replacing λ_p:

$$\mu_{nS}/\mu_n = 1-[(1-w_n)n_B\lambda_n/\Delta N]\times \times[\exp(\Delta_S u)-1],\ \Delta_S u \leqslant 0 \qquad (30a)$$

$$\mu_{nS}/\mu_n = 1-[(1-w_n)n_B\lambda_n/\Delta N]\left[\exp(\Delta_S u)-1- -(1-w_n)\int_{-\Delta_S u}^{u_B-u_a} dv\,\exp(-v+2K)/(1-w_n\exp 2K)\right],$$
$$\Delta_S u \geqslant 0 \qquad (30b)$$

4. ELLIPSOIDAL ENERGY SURFACES

Ellipsoidal energy surfaces were treated by HAM and MATTIS[20, 21]. For each valley a transformed k-space was introduced in which the valley becomes spherical and in which the Boltzmann equation boundary value treatment of SCHRIEFFER could be applied. (The revised treatment of Ref. 14 can also be applied.) The conductance so calculated is then transformed back to physical space where, added to the contributions of the other valleys, it forms a conductance tensor. For special orientations of the surface relative to the crystal axes the tensor is isotropic. Little work in this field has been published since SCHRIEFFER's 1956 review.[21]

5. MAGNETO-SURFACE EFFECTS

Surface mobility theory could be tested against experiment if an independent means, free of serious objections, were available for measuring u_S. This is not yet the case. Magneto-surface studies, begun by PETRITZ[22] and ZEMEL, circumvent this difficulty by comparing the theoretical and experimental relationships of two independent transport quantities, such as the change in surface conductivity versus surface magnetoresistance or Hall coefficient.

In PETRITZ's initial paper[22] the bulk region and the space-charge layer were treated as physically separated, but connected at the eletrodes. Bulk transport theory was applied to both regions so that the energy dependence of the bulk relaxation time could be considered. Constant mean free path, as in lattice scattering, was assumed. A model was used having a spherical conduction band and spherical heavy and light hole bands. Formulas for $R/R_{\max}$, $\Delta\rho/\rho H^2$, and $(\Delta\sigma-\Delta_m\sigma)/\Delta_m\sigma$ could then be developed from bulk theory and the circuit analog. From these, plots of $R/R_{\max}$ versus $(\Delta\sigma-\Delta_m\sigma)/\Delta_m\sigma$ and $\Delta\rho/\rho H^2$ versus $(\Delta\sigma-\Delta_m\sigma)/\Delta_m\sigma$ were made[19], for various light to heavy hole mobilities and superimposed on experimental plots made from measurements on intrinsic germanium. Rather good fit was obtained, for nearly flat bands, for heavy to light hole mobility ratios between eight and twelve, in agreement with WILLARDSON *et al.*[18]

To extend this analysis to the case $\Delta_S u > 1$ surface scattering corrections were made, by reducing the mobility of all the carriers in the same ratio as for the apparent hole mobility. It was recognized, however, that greater corrections[19] should be made for the light holes than for the heavy holes. Moreover, the Hall and magnetoresistivity mobilities were also reduced in the same ratio. The resultant plots of $R/R_{\max}$ and $\Delta\rho/\rho H^2$ versus $(\Delta\sigma-\Delta_m\sigma)/\Delta_m\sigma$ were then superimposed on those in which surface scattering was neglected. The former lie much closer to the experimental curves than do the latter. This provides strong evidence for the superiority of the diffuse scattering model to the specular reflection model.

The same effects were shown[23] on slightly p-type germanium by ZEMEL and PETRITZ. Figure 3 showing $R/R_{\max}$ versus $(\Delta\sigma-\Delta_m\sigma)/\Delta_m\sigma$ is taken from their work. The striking confirmation of the diffuse scattering model is clear. The agreement between experiment and theory is remarkable, considering the serious approximations in the latter.

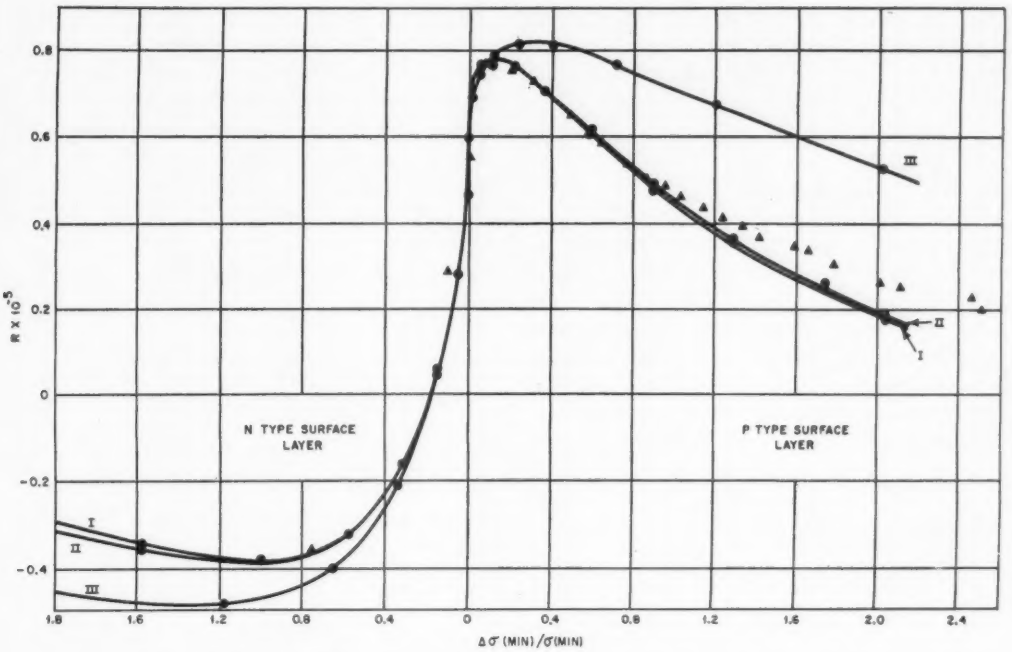


FIG. 3. Hall coefficient as a function of the fractional change in surface conductance with respect to the conductance minimum. The triangles refer to experiment, and the circles refer to theory. Curve (1) uses ZEMEL's expression.[24] Curve (2) uses the square of the conductivity effective mobility for $\langle \mu_S^2 \rangle / \langle \mu_B^2 \rangle$. Curve (3) has the effective mobilities set equal to the bulk mobilities (specular surface hypothesis). After ZEMEL and PETRITZ[23].

ZEMEL[24] has begun the Boltzmann equation treatment of magneto-surface effects by generalizing the SCHRIEFFER theory to include magnetic fields. He applied LAGRANGE's method again, and obtained a solution of the boundary value problem from which the galvanomagnetic coefficients were calculated. A scattering b.c. was applied, however, which suppresses the transport term of f at the surface for electrons moving both toward and away from the surface. This is stronger than the Fuchs b.c., equation (15), and probably overcorrects for surface scattering. ZEMEL finds a surface Hall mobility about 13 per cent less than the conductivity mobility for strong accumulation layers. AMITH[25] has used the revised boundary condition, as discussed in Section 1 above, corresponding to a finite bulk potential, and has calculated somewhat differently defined galvanomagnetic quantities.

Up to now mathematical complexity has prevented the use of any but a constant relaxation time τ in a Boltzmann equation treatment. This is far more serious for the galvanomagnetic effects than for the mobility, just as it is in the bulk.

6. THE SCATTERING BOUNDARY CONDITION

Although our knowledge of semiconductor surfaces is rather incomplete, it is worth while looking for the origin of the apparently successful FUCHS boundary condition, equation (29). It can be shown[26] that a b.c. of this form may be expected from a simple model of the surface, and that a value of $w \sim 0$ may be expected in germanium under ordinary circumstances.

The model consists of a specular surface containing a random array of N localized scatterers/cm^2, each having a differential scattering cross section $\sigma(\boldsymbol{k},\boldsymbol{k}')$. A b.c. can be derived by expressing the continuity of particle fluxes at the surface. To do this it is convenient to replace the physical surface by extending the semiconductor to fill all

space, and inserting a plane of reflection on which the scatterers lie. The fictitious half-space then has particle flux densities $\nu^{(2)}(\boldsymbol{k})$ which are related to those in the physical half-space by

$$\nu^{(2)}(\boldsymbol{k}) = \nu^{(1)}(\tilde{\boldsymbol{k}}) \tag{31}$$

where $\tilde{\boldsymbol{k}}$ is the reflection of $\boldsymbol{k}$.

A flat cylinder enclosing unit area of the surface is then envisioned. The flux density $\nu^{(1)}(\boldsymbol{k})$ incident from the physical space differs from the flux density $\nu^{(2)}(\boldsymbol{k})$ emerging in the same direction in the fictitious half-space because of particles scattered out of $\boldsymbol{k}$ into other directions $\boldsymbol{k}'$ and because of particles scattered from other directions $\boldsymbol{k}'$ into direction $\boldsymbol{k}$. Expressing this balance for the cylinder, applying equation (31) and the relations

$$\sigma(\boldsymbol{k}, \boldsymbol{k}') = \sigma(\boldsymbol{k}', \boldsymbol{k}) \tag{32}$$

$$\nu(\boldsymbol{k}) = f(\boldsymbol{k})v, \tag{33}$$

one finds the b.c.

$$f(\boldsymbol{k}) - f(\tilde{\boldsymbol{k}}) = (N\sigma_T \sec\theta) \left[f_1(\theta) - \sigma_T^{-1} \int d\Omega' \sigma(\boldsymbol{k},\boldsymbol{k}') f(\boldsymbol{k}') \right] \tag{34}$$

where θ is the angle between $\boldsymbol{k}$ and the normal to the surface and σ_T is the total scattering cross section of one center. This complicated form can be simplified if we recall that

$$f_1(\boldsymbol{k}) = \boldsymbol{E} \cdot \boldsymbol{k} g(k) \text{ for } N = 0 \tag{35}$$

i.e. in the specular case. Using this for the integral in the r.h.s. of equation (34) one finds[26] the FUCHS form

$$f_1(\boldsymbol{k}) - f_1(\tilde{\boldsymbol{k}}) = [1 - w(\theta)] f_1(\boldsymbol{k}) \tag{36}$$

where

$$w = 0 \text{ for } N\sigma_T\langle 1-\xi\rangle \geqslant 1 \tag{37a}$$

$$1 - w(\theta) = N\sigma_T\langle 1-\xi\rangle \sec\theta \text{ for } \cos\theta \geqslant N\sigma_T\langle 1-\xi\rangle \tag{37b}$$

$$w(\theta) = 0 \text{ for } \cos\theta \leqslant N\sigma_T\langle 1-\xi\rangle \tag{37c}$$

where $\langle\xi\rangle$ is the mean scattering angle for one scatterer.

If the individual centers are assumed to be Coulomb centers with Debye shielding, then one finds for thermal electrons at 300°K in germanium

$$\sigma_T\langle 1-\xi\rangle \approx 3 \times 10^{-12} \text{ cm}^2$$

so that at $N \sim 3 \times 10^{11}$ cm^{-2} the expected value of $w(\theta)$ is close to zero. It is interesting to recall that this is the order of magnitude of fast surface state densities in germanium.[27]

What this model assumes is that a clean, perfect germanium surface at low temperatures is specular. This would be suggested if such a surface has translational symmetry in directions parallel to the surface. For then exact propagating eigenfunctions should exist which carry surface currents. Such a surface would not be dissipative until perturbed by localized defects or by bulk or Rayleigh phonons.

REFERENCES

1. SHOCKLEY W. and PEARSON G. L., *Phys. Rev.* **74,** 232 (1948).
2. SCHOTTKY W., *Z. Phys.* **113,** 367 (1939).
   SCHOTTKY W., *Z. Phys.* **118,** 539 (1942).
   SCHOTTKY W. and SPENKE E., *Wiss. Veröff. Siemens* **18,** 3 (1939).
   MOTT N. F., *Proc. Roy. Soc.* A**171,** 27 (1939).
3. GARRETT C. G. B. and BRATTAIN W. H., *Phys. Rev.* **99,** 376 (1955).
4. KINGSTON R. H. and NEUSTADTER S., *J. Appl. Phys.* **26,** 718 (1955).
5. SEIWATZ R. and GREEN M., *J. Appl. Phys.* **29,** 1034 (1958).
6. SCHRIEFFER J. R., *Phys. Rev.* **97,** 641 (1955).
7. SHOCKLEY W., *Bell Syst. Tech. J.* **28,** 435 (1949).
8. FUCHS K., *Proc. Camb. Phil. Soc.* **34,** 100 (1938).
9. SONDHEIMER E. H., *Advanc. Phys.* **1,** 1 (1952).
10. ZEMEL J. N., *Bull. Amer. Phys. Soc.* (ser 2) **3,** 255 (1958).
11. FRANKL D. R., *Bull. Amer. Phys. Soc.* (ser 2) **4,** 179 (1959).
12. GREENE R. F., unpublished.
13. FLIETNER H., *Ann. Phys., Lpz.* (ser. 7) **3,** 396 (1959).
14. GREENE R. F., FRANKL D. R. and ZEMEL J. N., *Phys. Rev.* **118** (in press).
15. MILLER F. H., *Partial Differential Equations.* Wiley, New York (1941).
16. CONWELL E. M., *Proc. Inst. Radio Engrs. N.Y.* **46,** 1281 (1958).
17. LAX B. and MAVROIDES J. G., *Phys. Rev.* **100,** 1650 (1955).
18. WILLARDSON R. K., HARMAN T. C. and BEER A.C., *Phys. Rev.* **96,** 1512 (1954).
19. ZEMEL J. N. and PETRITZ R. L., *Phys. Rev.* **110,** 1263 (1958).
20. HAM F. S. and MATTIS D., *Univ. Illinois Tech. Rep.* No. 4 (1955), unpublished. (U.S. Army OOR, Contract DA-11-022-ORD-1001).
21. SCHRIEFFER J. R., *Semiconductor Surface Physics* (edited by KINGSTON R. H.) p. 55. University of Pennsylvania Press, Philadelphia (1957).
22. PETRITZ R. L., *Phys. Rev.* **110,** 1254 (1958).
23. ZEMEL J. N. and PETRITZ R. L., *J. Phys. Chem. Solids* **8,** 102 (1958).
24. ZEMEL J. N., *Phys. Rev.* **112,** 762 (1958).
25. A. AMITH, this Conference, preceding paper.
26. GREENE R. F., to be published.
27. MANY A., *J. Phys. Chem. Solids* **8,** 87 (1956).

VOL.
14
1960

DISCUSSION

A. MANY, N. B. GROVER, Y. GOLDSTEIN and E. HARNIK (The Hebrew University, Jerusalem, Israel): The calculations presented in this paper of μ_{nS}/μ_n as a function of u_S employ two basic parameters: the bulk potential u_B and the ratio of the length λ_n to the Debye length L_D, where L_D is given by $(\kappa kT/8\pi n_i e^2)^{1/2}$. This latter parameter is extremely temperature sensitive due to the exponential dependence of n_i on T. As such it is an inconvenient quantity whenever results of calculations at different temperatures are required. The purpose of this note is to call attention to the fact that in extrinsic samples ($u_B > \sim 3$ for n-type), both u_B and λ_n/L_D can be replaced by a single, relatively temperature-insensitive parameter λ_n/L, where L is the space-charge layer thickness defined by $(\kappa kT/4\pi e^2(N_D-N_A))^{1/2}$.

In this range the sample constitutes essentially a single carrier system, and the general Poisson equation for an n-type sample reduces to the simple form

$$L^2\partial^2(u-u_B)/\partial x^2 = e^{u-u_B}-1$$

to a very good approximation, provided that $-(u_S-u_B) < \sim (2u_B-3)$. Thus in terms of the variable $(u-u_B)$ the barrier shape is independent of u_B and is characterized by the thickness L. It is, therefore, the ratio of the mean free path ($\propto \lambda_n$) to *this* parameter which can be expected to determine the magnitude of the mobility reduction for any value of (u_S-u_B) in this range: The larger this ratio, the smaller is the μ_{nS}/μ_n expected.

With these physical considerations in mind, the theoretical results of GREENE *et al.** as reported in this paper were examined. It was found that for the extrinsic range under discussion equations (18) can be rewritten to within an accuracy of better than 1 per cent in the simpler forms

$$\mu_{nS}/\mu_n = 1-2^{-1/2}r(e^{\Delta_S u}-\Delta_S u-1)^{-1/2}(1-e^{\Delta_S u}),$$
$$u_B > \sim 3, \quad 3-2u_B \lesssim \Delta_S u \leqslant 0 \qquad (18c')$$

$$\mu_{nS}/\mu_n = 1-2^{-1/2}r(e^{\Delta_S u}-\Delta_S u-1)^{-1/2}\times$$
$$\times \int_{-\Delta_S u}^{0} dv e^{-v}(1-e^{-f(v)/r}),$$
$$u_B > \sim 3, \Delta_S u \geqslant 0 \qquad (18d')$$

where

$$f(v) = (2\pi)^{-1/2}\times$$
$$\times \int_{-v}^{\Delta_S u} dy(y+v)^{-1/2}(e^y-y-1)^{-1/2}, u_B > \sim 3 \qquad (18e')$$

and

$$\Delta_S u = u_S-u_B, \quad r = \lambda_n/L \propto \mu_n(N_D-N_A)^{1/2}$$

It is thus seen that μ_{nS}/μ_n is indeed independent of u_B and involves but a single parameter r whose only temperature dependence is that of the bulk mobility μ_n. An examination of the above expressions shows that for any value of $\Delta_S u$, the ratio μ_{nS}/μ_n decreases monotonically with increasing r, as deduced above from physical considerations.

The table below illustrates the range of variation of these parameters for an n-type sample ($N_D-N_A = 6{\cdot}6\times10^{13}$ cm^{-3}) at the three temperatures at which experimental results have been reported.†

T(°K)	u_B	λ_n/L_D	λ_n/L
250	4·4	$1{\cdot}1\times10^{-2}$	0·07
100	33·4	$2{\cdot}3\times10^{-8}$	0·30
80	44·7	$1{\cdot}2\times10^{-10}$	0·43

The advantages in using the single-carrier approximation are immediately apparent. It is seen that in this sample the requirement $u_B > \sim 3$ is amply fulfilled, and indeed this will be the case for most samples doped with group III or V elements for temperatures below room temperature. In place of the two parameters u_B and $\lambda_n L_D$ (the latter varying over 8 orders of magnitude), we have a single parameter λ_n/L whose temperature variation reflects directly that of the mobility reduction.

* R. F. GREENE, D. R. FRANKL and J. N. ZEMEL, private communication; also *Phys. Rev.*, to be published.

† A. MANY, N. B. GROVER, Y. GOLDSTEIN and E. HARNIK, this Conference, Fig. 4, p. 190.